```
                    GenCore version 5.1.7
           Copyright (c) 1993 - 2006  Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 20, 2006, 20:29:34 ; Search time 1060 Seconds
                                             (without alignments)
                                             4288.038 Million cell updates/sec

Title:          US-10-782-096-2
Perfect score:  3597
Sequence:       1 MNSYKNKNEYEMLDALRINS..........TFPNQSLEKREQEVNDLFIN 682

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters:      9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10782096/runat_20012006_095214_24386/app_query.fasta_1.839
-DB=N_Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10782096_@CGN_1_1_727_@runat_20012006_095214_24386 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_21:*
                1:  geneseqn1980s:*
                2:  geneseqn1990s:*
                3:  geneseqn2000s:*
                4:  geneseqn2001as:*
                5:  geneseqn2001bs:*
                6:  geneseqn2002as:*
                7:  geneseqn2002bs:*
                8:  geneseqn2003as:*
                9:  geneseqn2003bs:*
                10:  geneseqn2003cs:*
                11:  geneseqn2003ds:*
                12:  geneseqn2004as:*
                13:  geneseqn2004bs:*
                14:  geneseqn2005s:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3597	100.0	2049	13	ADR89407	Adr89407 AXMI-009
2	3536	98.3	2016	13	ADR89409	Adr89409 AXMI-009
3	3488	97.0	1986	13	ADR89411	Adr89411 AXMI-009
4	948	26.4	3621	6	ABK87234	Abk87234 Bacillus
5	948	26.4	3621	12	ADL15304	Adl15304 B thuring
6	948	26.4	3621	14	AEB45606	Aeb45606 B. thurin
7	948	26.4	3621	14	AEB34684	Aeb34684 Bacillus
8	948	26.4	4874	6	ABK87247	Abk87247 Bacillus
9	948	26.4	4874	12	ADL15320	Adl15320 B thuring
10	948	26.4	4874	14	AEB45622	Aeb45622 B. thurin
11	942.5	26.2	3507	2	AAT43221	Aat43221 Antiscara
12	940.5	26.1	3690	6	ABK51132	Abk51132 cDNA enco
13	926	25.7	2003	6	ABK87236	Abk87236 Bacillus
14	926	25.7	2010	6	ABK87241	Abk87241 Bacillus
15	926	25.7	2010	6	ABK87238	Abk87238 Maize opt
16	926	25.7	2010	12	ADL15308	Adl15308 B thuring
17	926	25.7	2010	12	ADL15314	Adl15314 B thuring
18	926	25.7	2010	14	AEB45610	Aeb45610 Maize-opt
19	926	25.7	2010	14	AEB45616	Aeb45616 B. thurin
20	925.5	25.7	2022	12	ADL15384	Adl15384 B thuring
21	925.5	25.7	2022	12	ADL15352	Adl15352 B thuring
22	925.5	25.7	2022	14	AEB45654	Aeb45654 B. thurin
23	925.5	25.7	2022	14	AEB45686	Aeb45686 B. thurin
24	925	25.7	2022	6	ABK87239	Abk87239 Bacillus
25	925	25.7	2022	12	ADL15336	Adl15336 B thuring
26	925	25.7	2022	12	ADL15310	Adl15310 B thuring
27	925	25.7	2022	12	ADL15328	Adl15328 B thuring
28	925	25.7	2022	12	ADL15332	Adl15332 B thuring
29	925	25.7	2022	12	ADL15372	Adl15372 B thuring
30	925	25.7	2022	14	AEB45638	Aeb45638 B. thurin
31	925	25.7	2022	14	AEB45612	Aeb45612 B. thurin
32	925	25.7	2022	14	AEB45634	Aeb45634 B. thurin
33	925	25.7	2022	14	AEB45630	Aeb45630 B. thurin
34	925	25.7	2022	14	AEB45674	Aeb45674 B. thurin
35	924.5	25.7	2022	12	ADL15346	Adl15346 B thuring
36	924.5	25.7	2022	12	ADL15378	Adl15378 B thuring
37	924.5	25.7	2022	14	AEB45648	Aeb45648 B. thurin
38	924.5	25.7	2022	14	AEB45680	Aeb45680 B. thurin
39	924	25.7	2025	12	ADL15348	Adl15348 B thuring
40	924	25.7	2025	12	ADL15380	Adl15380 B thuring
41	924	25.7	2025	14	AEB45682	Aeb45682 B. thurin
42	924	25.7	2025	14	AEB45650	Aeb45650 B. thurin
43	924	25.7	3633	6	ABK87235	Abk87235 Bacillus
44	924	25.7	3633	12	ADL15306	Adl15306 B thuring
45	924	25.7	3633	14	AEB45608	Aeb45608 B. thurin

ALIGNMENTS

```
RESULT 1
ADR89407
ID   ADR89407 standard; cDNA; 2049 BP.
XX
AC   ADR89407;
XX
DT   18-NOV-2004  (first entry)
XX
DE   AXMI-009 coding sequence.
XX
KW   ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW   expression cassette; transformation; transgenic; plant; bacteria;
KW   lepidoptera; coleoptera; pest; pesticide; resistance;
KW   pesticidal activity.
XX
OS   Bacillus thuringiensis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2049
FT                   /*tag=  a
FT                   /product= "AXMI-009"
XX
PN   WO2004074462-A2.
XX
PD   02-SEP-2004.
XX
PF   20-FEB-2004; 2004WO-US005829.
XX
```

```
PR   20-FEB-2003; 2003US-0448632P.
PR   20-FEB-2003; 2003US-0448633P.
PR   20-FEB-2003; 2003US-0448797P.
PR   20-FEB-2003; 2003US-0448806P.
PR   20-FEB-2003; 2003US-0448810P.
PR   20-FEB-2003; 2003US-0448812P.
PR   19-FEB-2004; 2004US-00781979.
PR   19-FEB-2004; 2004US-00782020.
PR   19-FEB-2004; 2004US-00782096.
PR   19-FEB-2004; 2004US-00782141.
PR   19-FEB-2004; 2004US-00782570.
PR   19-FEB-2004; 2004US-00783417.
XX
PA   (ATHE-) ATHENIX CORP.
XX
PI   Carozzi N,  Hargiss T,  Koziel MG,  Duck NB,  Carr B;
XX
DR   WPI; 2004-635574/61.
DR   P-PSDB; ADR89408.
XX
PT   New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT   and polypeptides, useful for killing lepidopteran or coleopteran pests or
PT   for producing organisms with pesticide resistance.
XX
PS   Claim 1; SEQ ID NO 19; 178pp; English.
XX
CC   This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC   endotoxin coding sequences of the invention have alternative start
CC   codons, producing more than one protein from a single open reading frame.
CC   The nucleic acid sequences of the invention are useful in DNA constructs
CC   or expression cassettes for transformation and expression in plants and
CC   bacteria. The nucleic acids and corresponding polypeptides are useful for
CC   killing lepidopteran or coleopteran pests. Compositions containing the
CC   delta-endotoxins of the invention, and methods for their production, are
CC   useful for the production of organisms with pesticide resistance,
CC   specifically bacteria and plants. These organisms are useful for
CC   generating altered or improved delta-endotoxin or delta-endotoxin-
CC   associated proteins that have pesticidal activity, or for detecting the
CC   presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC   nucleic acids in products or organisms.
XX
SQ   Sequence 2049 BP; 688 A; 316 C; 403 G; 642 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0           Length:         2049
Score:                  3597.00     Matches:        682
Percent Similarity:     100.0%      Conservative:   0
Best Local Similarity:  100.0%      Mismatches:     0
Query Match:            100.0%      Indels:         0
DB:                     13          Gaps:           0

US-10-782-096-2 (1-682) x ADR89407 (1-2049)

Qy          1 MetAsnSerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATTCATATAAAAATAAAAATGAATATGAAATGTTGGATGCTTTACGAATCAACTCT 60

Qy         21 AsnMetSerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATATGTCTAATTGTTATCCAAGGTATCCACTAGCAAAAGATCCACAAATGACTATGCGA 120

Qy         41 AsnThrAsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAsp 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACACGAACTATAAAGAATGGCTAAATATGTGTGATTCAAATACACAATTTATTGGTGAT 180

Qy         61 IleSerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAAGCACGTATTCTAGCCCTGAAGCTGCTTTAAGTGTACGAGATGCTGTTTTAACGGGT 240

Qy         81 IleAsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTAACAGTGTAGGGACTATACTTTCGAATTTAGGGGTCCCTTTGGCAAGTCAATCATTT 300

Qy        101 GlyIleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAspProPheGluAlaLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAATAATTAGTAGGCTAATAGGTATTTTATGGGCAGGGCCTGATCCATTTGAAGCACTT 360

Qy        121 MetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGTTCTTGTTGAAGAGCTTATTAAGAAAAGTATAGATCAGCGTGTAAGAGAAAATGCT 420

Qy        141 LeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAla 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTAGAGAGCTAGAAGGTTTACAGGGAATTATGAGACTATATCAAACTAGACTGCAAGCA 480

Qy        161 TrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThrGlnTyrAlaIleVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGCTAGTTAACAAGAATGATGACAATCGGAGGGCACTAGTAACGCAGTATGCAATTGTT 540

Qy        181 AspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATAACTTTTTCGAAAAGAATATGCCAAAATTCAAGGAAAGAAACTTTGAAATTTTATTG 600

Qy        201 LeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTACCAGTATATGCACAAGCCGCGAATTTGCATTTAATTTTATTAAGAGATGCTGATTAT 660

Qy        221 PheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTGGAGCACAGTGGCAATTAGGTGATGATGAAATTCGTGATAATTATATCAGACTACAA 720

Qy        241 GlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGln 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACTGATTAGAGAATATAAAGATCATTGTATAACATTCTATAACCAGGGTTTAAATCAA 780

Qy        261 PheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTAATCGCTCAAATGCTCAAGATTGGGTGAGCTTTAATAGGTTTCGTACAGATATGACA 840

Qy        281 LeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTAACAGTATTAGATCTCGCAATATTATTTCCAAACTATGATCCACGTAGGTATCCATTA 900

Qy        301 AlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCAGTAAAAACGGAATTGACTAGGGAAGTTTATACAGATCCAGTAGGGTTTACTGGGGTA 960

Qy        321 LeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAla 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAGAAAGTGGAGGTAGGACTTACCCTTGGTATAATCCTAATAATACAACCTTTACTGCT 1020

Qy        341 MetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheVal 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATGGAAAATAACGCAAGACGACGTCCTTCTTATACCACTTGGCTTAATCGTATTTTTGTA 1080

Qy        361 TyrThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHisThrLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TATACAAGGACTCTAGGTAATATGTCTGATGTGAGAAATATTTGGGGAGGGCATACATTA 1140

Qy        381 ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAspSerIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTTGAAAATGGAAATGATGGTTCTGAAATAACCCATAACTTTGGTAAAACTGATTCTATT 1200

Qy        401 ThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAla 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCCTATTCAATATTTTAATTTCGCGAACCTTTCTGTTTTCAGTATTGAGTCACTTGCT 1260

Qy        421 ArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGTATATATTTAGGAGGAACAGAGGCTAATAATTATATTACTAGTCAGTATGGAGTCTCG 1320

Qy        441 ArgValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluVal 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAGTTATTTTTAATACATCAAATATAAATAATGTACCTGGATCTTTAAGATACGAAGTG 1380

Qy        461 ProAlaAsnLeuProSerGlnThrIleLeuSerGluLeuProGlyLysAspLysProArg 480
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGCTAATCTTCCATCCCAAACTATATTATCAGAATTACCAGGAAAGGATAAGCCAAGA 1440

Qy        481 ProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArg 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCAAACGCAGGAGATTTCAGCCATAGATTATCTTATATATCAAATTTTGATGCACGGCGA 1500

Qy        501 SerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAsp 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGTAGTTCAGGCGGTATTGTTAGTCTTTTAACGTTTGGTTGGGCACATACCAGTATGGAT 1560

Qy        521 ArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CGTAATAATCGTCTTGAACCAGATAAAATTACTCAAATAGATGCAGTTAAAGGTTGGGGG 1620

Qy        541 GlyAsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAsp 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGAATATCGGGTTTGTCATCCCAGGACCTACTGGGGGGAATTTGGTAAAAGTCAGTGAT 1680

Qy        561 SerTrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeu 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGTTGGCATTCACTTAAAGTTCAAGCACCACAAAGACAAACAAGTTATCGTATTCGTTTG 1740

Qy        581 ArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSer 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGTTATGCTTGTTTAGTTACCCATGGGGATGCTATTTTTGTAGAACACAGCGGCAGTAGT 1800

Qy        601 HisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeu 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CATATAGTTTCATTTTTTGATTGCTCAAATTCATCAGGTCGTCCATCAAACACTCTTCTA 1860

Qy        621 GluSerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GAGAGTGATTTTCGCTATATTGATGTTCCAGGTATTTTTACACCATCAATAAATCCCTTA 1920

Qy        641 IleArgTyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIlePro 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATAAGATATAGAACACAAAGCTTTGGTACCCACGCGATAGACAAATTTGAATTTATTCCA 1980

Qy        661 LeuAsnThrPheProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPhe 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CTTAACACTTTTCCGAATCAATCATTAGAAAAAAGAGAACAGGAAGTAAATGATCTATTT 2040

Qy        681 IleAsn 682
              ||||||
Db       2041 ATCAAT 2046
```

```
RESULT 2
ADR89409
ID   ADR89409 standard; cDNA; 2016 BP.
XX
AC   ADR89409;
XX
DT   18-NOV-2004  (first entry)
XX
DE   AXMI-009 alternative start site coding sequence.
XX
KW   ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW   expression cassette; transformation; transgenic; plant; bacteria;
KW   lepidoptera; coleoptera; pest; pesticide; resistance;
KW   pesticidal activity.
XX
OS   Bacillus thuringiensis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2016
FT                   /*tag=  a
FT                   /product= "Alternative AXMI-009"
XX
PN   WO2004074462-A2.
XX
PD   02-SEP-2004.
XX
PF   20-FEB-2004; 2004WO-US005829.
XX
PR   20-FEB-2003; 2003US-0448632P.
PR   20-FEB-2003; 2003US-0448633P.
PR   20-FEB-2003; 2003US-0448797P.
PR   20-FEB-2003; 2003US-0448806P.
PR   20-FEB-2003; 2003US-0448810P.
PR   20-FEB-2003; 2003US-0448812P.
PR   19-FEB-2004; 2004US-00781979.
PR   19-FEB-2004; 2004US-00782020.
PR   19-FEB-2004; 2004US-00782096.
PR   19-FEB-2004; 2004US-00782141.
PR   19-FEB-2004; 2004US-00782570.
PR   19-FEB-2004; 2004US-00783417.
XX
PA   (ATHE-) ATHENIX CORP.
XX
PI   Carozzi N,  Hargiss T,  Koziel MG,  Duck NB,  Carr B;
XX
DR   WPI; 2004-635574/61.
DR   P-PSDB; ADR89410.
XX
PT   New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT   and polypeptides, useful for killing lepidopteran or coleopteran pests or
PT   for producing organisms with pesticide resistance.
XX
PS   Claim 1; SEQ ID NO 21; 178pp; English.
XX
CC   This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC   endotoxin coding sequences of the invention have alternative start
CC   codons, producing more than one protein from a single open reading frame.
CC   The nucleic acid sequences of the invention are useful in DNA constructs
CC   or expression cassettes for transformation and expression in plants and
CC   bacteria. The nucleic acids and corresponding polypeptides are useful for
CC   killing lepidopteran or coleopteran pests. Compositions containing the
CC   delta-endotoxins of the invention, and methods for their production, are
CC   useful for the production of organisms with pesticide resistance,
CC   specifically bacteria and plants. These organisms are useful for
CC   generating altered or improved delta-endotoxin or delta-endotoxin-
CC   associated proteins that have pesticidal activity, or for detecting the
CC   presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC   nucleic acids in products or organisms.
XX
SQ   Sequence 2016 BP; 668 A; 315 C; 400 G; 633 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               0           Length:        2016
Score:                   3536.00     Matches:       671
Percent Similarity:      100.0%      Conservative:  0
Best Local Similarity:   100.0%      Mismatches:    0
Query Match:             98.3%       Indels:        0
DB:                      13          Gaps:          0

US-10-782-096-2 (1-682) x ADR89409 (1-2016)

Qy         12 MetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyrProArgTyrProLeu 31
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTTGGATGCTTTACGAATCAACTCTAATATGTCTAATTGTTATCCAAGGTATCCACTA 60

Qy         32 AlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGluTrpLeuAsnMetCys 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAAAAGATCCACAAATGACTATGCGAAACACGAACTATAAAGAATGGCTAAATATGTGT 120

Qy         52 AspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSerProGluAlaAlaLeu 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATTCAAATACACAATTTATTGGTGATATAAGCACGTATTCTAGCCCTGAAGCTGCTTTA 180

Qy         72 SerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThrIleLeuSerAsnLeu 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTGTACGAGATGCTGTTTTAACGGGTATTAACAGTGTAGGGACTATACTTTCGAATTTA 240

Qy         92 GlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeuIleGlyIleLeuTrp 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 GGGGTCCCTTTGGCAAGTCAATCATTTGGAATAATTAGTAGGCTAATAGGTATTTTATGG 300

Qy        112 AlaGlyProAspProPheGluAlaLeuMetValLeuValGluGluLeuIleLysLysSer 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCAGGGCCTGATCCATTTGAAGCACTTATGGTTCTTGTTGAAGAGCTTATTAAGAAAAGT 360

Qy        132 IleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGluGlyLeuGlnGlyIleMet 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATAGATCAGCGTGTAAGAGAAAATGCTCTTAGAGAGCTAGAAGGTTTACAGGGAATTATG 420

Qy        152 ArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLysAsnAspAspAsnArgArg 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGACTATATCAAACTAGACTGCAAGCATGGCTAGTTAACAAGAATGATGACAATCGGAGG 480

Qy        172 AlaLeuValThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPhe 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACTAGTAACGCAGTATGCAATTGTTGATAACTTTTTCGAAAAGAATATGCCAAAATTC 540

Qy        192 LysGluArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHis 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGGAAAGAAACTTTGAAATTTTATTGTTACCAGTATATGCACAAGCCGCGAATTTGCAT 600

Qy        212 LeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGlu 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTAATTTTATTAAGAGATGCTGATTATTTTGGAGCACAGTGGCAATTAGGTGATGATGAA 660

Qy        232 IleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIle 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTCGTGATAATTATATCAGACTACAAGGACTGATTAGAGAATATAAAGATCATTGTATA 720

Qy        252 ThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSer 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACATTCTATAACCAGGGTTTAAATCAATTTAATCGCTCAAATGCTCAAGATTGGGTGAGC 780

Qy        272 PheAsnArgPheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPhePro 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTAATAGGTTTCGTACAGATATGACATTAACAGTATTAGATCTCGCAATATTATTTCCA 840

Qy        292 AsnTyrAspProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyr 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACTATGATCCACGTAGGTATCCATTAGCAGTAAAAACGGAATTGACTAGGGAAGTTTAT 900

Qy        312 ThrAspProValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyr 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACAGATCCAGTAGGGTTTACTGGGGTATTAGAAAGTGGAGGTAGGACTTACCCTTGGTAT 960

Qy        332 AsnProAsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyr 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATCCTAATAATACAACCTTTACTGCTATGGAAAATAACGCAAGACGACGTCCTTCTTAT 1020

Qy        352 ThrThrTrpLeuAsnArgIlePheValTyrThrArgThrLeuGlyAsnMetSerAspVal 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACCACTTGGCTTAATCGTATTTTTGTATATACAAGGACTCTAGGTAATATGTCTGATGTG 1080

Qy        372 ArgAsnIleTrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThr 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAAATATTTGGGGAGGGCATACATTAGTTGAAAATGGAAATGATGGTTCTGAAATAACC 1140

Qy        392 HisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeu 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CATAACTTTGGTAAAACTGATTCTATTACTCCTATTCAATATTTTAATTTCGCGAACCTT 1200

Qy        412 SerValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsn 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTGTTTTCAGTATTGAGTCACTTGCTCGTATATATTTAGGAGGAACAGAGGCTAATAAT 1260

Qy        432 TyrIleThrSerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsn 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TATATTACTAGTCAGTATGGAGTCTCGAGAGTTATTTTTAATACATCAAATATAAATAAT 1320

Qy        452 ValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeuSer 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTACCTGGATCTTTAAGATACGAAGTGCCTGCTAATCTTCCATCCCAAACTATATTATCA 1380

Qy        472 GluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSer 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GAATTACCAGGAAAGGATAAGCCAAGACCAAACGCAGGAGATTTCAGCCATAGATTATCT 1440

Qy        492 TyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThr 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TATATATCAAATTTTGATGCACGGCGAAGTAGTTCAGGCGGTATTGTTAGTCTTTTAACG 1500

Qy        512 PheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThr 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTTGGTTGGGCACATACCAGTATGGATCGTAATAATCGTCTTGAACCAGATAAAATTACT 1560

Qy        532 GlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPheValIleProGlyProThr 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CAAATAGATGCAGTTAAAGGTTGGGGGGGGAATATCGGGTTTGTCATCCCAGGACCTACT 1620

Qy        552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGGGGAATTTGGTAAAAGTCAGTGATAGTTGGCATTCACTTAAAGTTCAAGCACCACAA 1680

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACAAACAAGTTATCGTATTCGTTTGCGTTATGCTTGTTTAGTTACCCATGGGGATGCT 1740

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATTTTTGTAGAACACAGCGGCAGTAGTCATATAGTTTCATTTTTTGATTGCTCAAATTCA 1800

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCAGGTCGTCCATCAAACACTCTTCTAGAGAGTGATTTTCGCTATATTGATGTTCCAGGT 1860

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArgThrGlnSerPheGlyThrHis 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ATTTTTACACCATCAATAAATCCCTTAATAAGATATAGAACACAAAGCTTTGGTACCCAC 1920

Qy        652 AlaIleAspLysPheGluPheIleProLeuAsnThrPheProAsnGlnSerLeuGluLys 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCGATAGACAAATTTGAATTTATTCCACTTAACACTTTTCCGAATCAATCATTAGAAAAA 1980

Qy        672 ArgGluGlnGluValAsnAspLeuPheIleAsn 682
              |||||||||||||||||||||||||||||||||
Db       1981 AGAGAACAGGAAGTAAATGATCTATTTATCAAT 2013
```

```
RESULT 3
ADR89411
ID   ADR89411 standard; cDNA; 1986 BP.
XX
AC   ADR89411;
XX
DT   18-NOV-2004  (first entry)
XX
DE   AXMI-009 alternative start site coding sequence #2.
XX
KW   ss; gene; delta-endotoxin; delta-endotoxin associate polypeptide;
KW   expression cassette; transformation; transgenic; plant; bacteria;
KW   lepidoptera; coleoptera; pest; pesticide; resistance;
KW   pesticidal activity.
XX
OS   Bacillus thuringiensis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1986
FT                   /*tag=  a
FT                   /product= "Alternative AXMI-009 #2"
XX
PN   WO2004074462-A2.
XX
PD   02-SEP-2004.
XX
PF   20-FEB-2004; 2004WO-US005829.
XX
```

```
PR   20-FEB-2003; 2003US-0448632P.
PR   20-FEB-2003; 2003US-0448633P.
PR   20-FEB-2003; 2003US-0448797P.
PR   20-FEB-2003; 2003US-0448806P.
PR   20-FEB-2003; 2003US-0448810P.
PR   20-FEB-2003; 2003US-0448812P.
PR   19-FEB-2004; 2004US-00781979.
PR   19-FEB-2004; 2004US-00782020.
PR   19-FEB-2004; 2004US-00782096.
PR   19-FEB-2004; 2004US-00782141.
PR   19-FEB-2004; 2004US-00782570.
PR   19-FEB-2004; 2004US-00783417.
XX
PA   (ATHE-) ATHENIX CORP.
XX
PI   Carozzi N,  Hargiss T,  Koziel MG,  Duck NB,  Carr B;
XX
DR   WPI; 2004-635574/61.
DR   P-PSDB; ADR89412.
XX
PT   New isolated delta-endotoxin and delta-endotoxin-associated nucleic acids
PT   and polypeptides, useful for killing lepidopteran or coleopteran pests or
PT   for producing organisms with pesticide resistance.
XX
PS   Claim 1; SEQ ID NO 23; 178pp; English.
XX
CC   This sequence encodes an isolated delta-endotoxin. Some of the delta-
CC   endotoxin coding sequences of the invention have alternative start
CC   codons, producing more than one protein from a single open reading frame.
CC   The nucleic acid sequences of the invention are useful in DNA constructs
CC   or expression cassettes for transformation and expression in plants and
CC   bacteria. The nucleic acids and corresponding polypeptides are useful for
CC   killing lepidopteran or coleopteran pests. Compositions containing the
CC   delta-endotoxins of the invention, and methods for their production, are
CC   useful for the production of organisms with pesticide resistance,
CC   specifically bacteria and plants. These organisms are useful for
CC   generating altered or improved delta-endotoxin or delta-endotoxin-
CC   associated proteins that have pesticidal activity, or for detecting the
CC   presence of delta-endotoxin or delta-endotoxin-associated proteins or
CC   nucleic acids in products or organisms.
XX
SQ   Sequence 1986 BP; 659 A; 310 C; 395 G; 622 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.76e-313      Length:         1986
Score:                  3488.00        Matches:        661
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            97.0%          Indels:         0
DB:                     13             Gaps:           0

US-10-782-096-2 (1-682) x ADR89411 (1-1986)

Qy         22 MetSerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsn 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCTAATTGTTATCCAAGGTATCCACTAGCAAAAGATCCACAAATGACTATGCGAAAC 60

Qy         42 ThrAsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIle 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACGAACTATAAAGAATGGCTAAATATGTGTGATTCAAATACACAATTTATTGGTGATATA 120

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCACGTATTCTAGCCCTGAAGCTGCTTTAAGTGTACGAGATGCTGTTTTAACGGGTATT 180

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGTGTAGGGACTATACTTTCGAATTTAGGGGTCCCTTTGGCAAGTCAATCATTTGGA 240

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAspProPheGluAlaLeuMet 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATAATTAGTAGGCTAATAGGTATTTTATGGGCAGGGCCTGATCCATTTGAAGCACTTATG 300

Qy        122 ValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeu 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTCTTGTTGAAGAGCTTATTAAGAAAAGTATAGATCAGCGTGTAAGAGAAAATGCTCTT 360

Qy        142 ArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrp 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAGAGCTAGAAGGTTTACAGGGAATTATGAGACTATATCAAACTAGACTGCAAGCATGG 420

Qy        162 LeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThrGlnTyrAlaIleValAsp 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTAGTTAACAAGAATGATGACAATCGGAGGGCACTAGTAACGCAGTATGCAATTGTTGAT 480

Qy        182 AsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeu 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACTTTTTCGAAAAGAATATGCCAAAATTCAAGGAAAGAAACTTTGAAATTTTATTGTTA 540

Qy        202 ProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPhe 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCAGTATATGCACAAGCCGCGAATTTGCATTTAATTTTATTAAGAGATGCTGATTATTTT 600

Qy        222 GlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGly 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGCACAGTGGCAATTAGGTGATGATGAAATTCGTGATAATTATATCAGACTACAAGGA 660

Qy        242 LeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPhe 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGATTAGAGAATATAAAGATCATTGTATAACATTCTATAACCAGGGTTTAAATCAATTT 720

Qy        262 AsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeu 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATCGCTCAAATGCTCAAGATTGGGTGAGCTTTAATAGGTTTCGTACAGATATGACATTA 780

Qy        282 ThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAla 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACAGTATTAGATCTCGCAATATTATTTCCAAACTATGATCCACGTAGGTATCCATTAGCA 840

Qy        302 ValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeu 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTAAAAACGGAATTGACTAGGGAAGTTTATACAGATCCAGTAGGGTTTACTGGGGTATTA 900

Qy        322 GluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMet 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAAGTGGAGGTAGGACTTACCCTTGGTATAATCCTAATAATACAACCTTTACTGCTATG 960

Qy        342 GluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyr 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAAAATAACGCAAGACGACGTCCTTCTTATACCACTTGGCTTAATCGTATTTTTGTATAT 1020

Qy        362 ThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHisThrLeuVal 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAAGGACTCTAGGTAATATGTCTGATGTGAGAAATATTTGGGGAGGGCATACATTAGTT 1080

Qy        382 GluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAspSerIleThr 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GAAAATGGAAATGATGGTTCTGAAATAACCCATAACTTTGGTAAAACTGATTCTATTACT 1140

Qy        402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCTATTCAATATTTTAATTTCGCGAACCTTTCTGTTTTCAGTATTGAGTCACTTGCTCGT 1200

Qy        422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATATATTTAGGAGGAACAGAGGCTAATAATTATATTACTAGTCAGTATGGAGTCTCGAGA 1260

Qy        442 ValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValPro 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTATTTTTAATACATCAAATATAAATAATGTACCTGGATCTTTAAGATACGAAGTGCCT 1320

Qy        462 AlaAsnLeuProSerGlnThrIleLeuSerGluLeuProGlyLysAspLysProArgPro 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCTAATCTTCCATCCCAAACTATATTATCAGAATTACCAGGAAAGGATAAGCCAAGACCA 1380

Qy        482 AsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSer 501
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 AACGCAGGAGATTTCAGCCATAGATTATCTTATATATCAAATTTTGATGCACGGCGAAGT 1440

Qy       502 SerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArg 521
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 AGTTCAGGCGGTATTGTTAGTCTTTTAACGTTTGGTTGGGCACATACCAGTATGGATCGT 1500

Qy       522 AsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGly 541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 AATAATCGTCTTGAACCAGATAAAATTACTCAAATAGATGCAGTTAAAGGTTGGGGGGGG 1560

Qy       542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 AATATCGGGTTTGTCATCCCAGGACCTACTGGGGGAATTTGGTAAAAGTCAGTGATAGT 1620

Qy       562 TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArg 581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 TGGCATTCACTTAAAGTTCAAGCACCACAAAGACAAACAAGTTATCGTATTCGTTTGCGT 1680

Qy       582 TyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHis 601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 TATGCTTGTTTAGTTACCCATGGGGATGCTATTTTTGTAGAACACAGCGGCAGTAGTCAT 1740

Qy       602 IleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGlu 621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 ATAGTTTCATTTTTTGATTGCTCAAATTCATCAGGTCGTCCATCAAACACTCTTCTAGAG 1800

Qy       622 SerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeuIle 641
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 AGTGATTTTCGCTATATTGATGTTCCAGGTATTTTTACACCATCAATAAATCCCTTAATA 1860

Qy       642 ArgTyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeu 661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 AGATATAGAACACAAAGCTTTGGTACCCACGCGATAGACAAATTTGAATTTATTCCACTT 1920

Qy       662 AsnThrPheProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 AACACTTTTCCGAATCAATCATTAGAAAAAAGAGAACAGGAAGTAAATGATCTATTTATC 1980

Qy       682 Asn 682
             |||
Db      1981 AAT 1983
```

RESULT 4
ABK87234

```
ID   ABK87234 standard; DNA; 3621 BP.
XX
AC   ABK87234;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Bacillus thuringiensis Cry1218-1 gene sequence.
XX
KW   Pesticidal; spraying; dusting; broadcasting; seed coating; insect pest;
KW   Colorado potato beetle; western corn rootworm; southern corn rootworm;
KW   insect target range; endotoxin; Cry1218; gene; ds.
XX
OS   Bacillus thuringiensis.
XX
PN   WO200234774-A2.
XX
PD   02-MAY-2002.
XX
PF   24-OCT-2001; 2001WO-US045468.
XX
PR   24-OCT-2000; 2000US-0242838P.
PR   23-OCT-2001; 2001US-00032717.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad AR,  Duck NB,  Feng X,  Flannagan RD,  Kahn TW,  Sims LE;
XX
DR   WPI; 2002-519178/55.
DR   P-PSDB; AAU99255.
XX
PT   New isolated pesticidal polypeptide useful for impacting insect pest e.g.
PT   Colorado potato beetle.
XX
PS   Claim 1; Page 91-96; 176pp; English.
XX
CC   The present invention relates to a new pesticidal polypeptide. The
CC   invention is useful for impacting an insect pest by applying the the
CC   molecules of the invention to the environment of the insect pest by
CC   spraying, dusting, broadcasting, or seed coating, where the insect pest
CC   is selected from Colorado potato beetle, western corn rootworm or
CC   southern corn rootworm. The invention is also useful for increasing
CC   insect target range and for producing transgenic microorganisms and
CC   plants that express the pesticidal polypeptide. The invention is also
CC   useful for producing transformed plants and in transforming any organism
CC   to produce the pesticidal polypeptide of the invention. The present
CC   nucleic acid sequence encodes a Bacillus thuringiensis wild-type Cry1218
CC   endotoxin protein
XX
SQ   Sequence 3621 BP; 1277 A; 608 C; 771 G; 965 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.61e-77        Length:         3621
Score:                  948.00          Matches:        248
Percent Similarity:     50.9%           Conservative:   119
Best Local Similarity:  34.4%           Mismatches:     276
Query Match:            26.4%           Indels:         78
DB:                     6               Gaps:           25

US-10-782-096-2 (1-682) x ABK87234 (1-3621)

Qy         3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
             |||         |||:::||||||||||||::::::||||||      ::::::::::::
Db         4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy        23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
             ||||||         |||||||||   |||   :::|||         ::::::|||   
Db        61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy        43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
             :::||||||::::::|||   |||                        |||:::   |||
Db       121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy        63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                |||      |||||||||      :::|||   :::||||||            |||
Db       163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy        82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
             :::   ||||||   :::||||||   ||||||||||||                     
Db       223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy       102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
             :::   ::::::||||||   |||||||||         :::         :::|||   
Db       283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy       120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                |||      |||||||||||||||   :::   |||   :::      |||      
Db       343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy       140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
             ||||||   |||||||||||||||            :::||||||   |||   |||:::
Db       403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy       160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                |||      |||   |||               |||   :::      ::::::   
Db       463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy       179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
             |||:::|||:::   |||   :::   ||||||   |||:::      |||||||||:::
Db       523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy       199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
```

```
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238

Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258

Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278

Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298

Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318

Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338

Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358

Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376

Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395

Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415

Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430

Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449

Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469

Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482

Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502

Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522

Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542

Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558

Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571

Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591

Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611

Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631

Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644

Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662

Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681

Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682

Db       2053 AAT 2055
```

```
RESULT 5
ADL15304
ID   ADL15304 standard; DNA; 3621 BP.
XX
AC   ADL15304;
XX
DT   06-MAY-2004  (first entry)
XX
DE   B thuringiensis strain 1218 Cry1218-1 (Cry8Bb1) DNA.
XX
KW   endotoxin; proteolytic site; insect gut; pesticidal; agriculture;
KW   plant pest; Cry8-like delta-endotoxin; Cry1218-1; Cry8Bb1; ds; gene.
XX
OS   Bacillus thuringiensis.
XX
PN   WO2004003148-A2.
XX
PD   08-JAN-2004.
XX
PF   25-JUN-2003; 2003WO-US020027.
XX
PR   26-JUN-2002; 2002US-0391786P.
PR   04-APR-2003; 2003US-0460787P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad AR,  Flannagan RD,  Herrmann R,  Kahn TW,  Lu AL,  Mccutchen BF;
PI   Presnail JK,  Wong JFH,  Yu C;
XX
DR   WPI; 2004-083033/08.
DR   P-PSDB; ADL15305.
XX
PT   New endotoxin comprising a mutation consisting of the alteration of at
PT   least one proteolytic site, where stability of the endotoxin in insect
PT   gut is increased, useful for providing improved pesticidal activity.
XX
PS   Example 4; SEQ ID NO 1; 333pp; English.
XX
CC   The invention relates to a novel endotoxin having a mutation consisting
```

```
CC   of the alteration of at least one proteolytic site, where the stability
CC   of the endotoxin in an insect gut is increased relative to an endotoxin
CC   lacking the mutation. The endotoxin of the invention may be useful for
CC   providing improved pesticidal activity against insect pests whilst the
CC   compositions of the invention may be useful in agriculture for
CC   controlling pests, including plant pests. The current sequence is that of
CC   the a DNA related to Bacillus thuringiensis strain 1218 Cry8-like delta-
CC   endotoxin Cry1218 of the invention.
XX
SQ   Sequence 3621 BP; 1277 A; 608 C; 771 G; 965 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.61e-77       Length:         3621
Score:                  948.00         Matches:        248
Percent Similarity:     50.9%          Conservative:   119
Best Local Similarity:  34.4%          Mismatches:     276
Query Match:            26.4%          Indels:         78
DB:                     12             Gaps:           25

US-10-782-096-2 (1-682) x ADL15304 (1-3621)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db        4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db       61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db      283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db      343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db      403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db      463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db      523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db      583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db      643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db      703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db      763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db      823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db      883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db      943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db      988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db     1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db     1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db     1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy      416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db     1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db     1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db     1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy      470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db     1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db     1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db     1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db     1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db     1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy      559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db     1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db     1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy      592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
```

```
                    |||    |||          :::    |||
Db        1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG--------------ATGCCAAAA 1812

Qy         612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
               :::        |||          |||          |||:::      |||
Db        1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy         632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                        :::   :::      :::     :::
Db        1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy         645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                  :::          |||          :::|||:::   |||||||||||::::::
Db        1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy         663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
               |||:::          |||   ||||||        ::::::   ||||||   ||||||
Db        1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy         682 Asn 682
               |||
Db        2053 AAT 2055
```

RESULT 6

AEB45606

```
ID   AEB45606 standard; DNA; 3621 BP.
XX
AC   AEB45606;
XX
DT   22-SEP-2005  (first entry)
XX
DE   B. thuringiensis Cry toxin, Cry1218-1 DNA.
XX
KW   Toxin; ds; gene; transgenic plant; insecticide; pesticide;
KW   plant insect pest; Cry8.
XX
OS   Bacillus thuringiensis; strain 1218-1.
XX
PN   WO2005066349-A2.
XX
PD   21-JUL-2005.
XX
PF   09-DEC-2004; 2004WO-US041530.
XX
PR   24-DEC-2003; 2003US-00746914.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad A,  Dong H,  Herrmann R,  Lu A,  Mccutchen BF,  Rice JA;
PI   Schepers EJ,  Wong JF;
XX
DR   WPI; 2005-533612/54.
DR   P-PSDB; AEB45607.
XX
PT   Novel isolated nucleic acid molecule having nucleotide sequence encoding
PT   pesticidal polypeptide comprising engineered proteolytic protection site
PT   resistant to protease, useful for protecting plant from pest e.g.
PT   Colorado potato beetle.
XX
PS   Example 4; SEQ ID NO 1; 362pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule comprising a
CC   nucleotide sequence encoding a pesticidal polypeptide having an
CC   engineered proteolytic protection site, which is not sensitive to the
CC   plant protease and protects the pesticidal polypeptide from proteolytic
CC   inactivation in a plant. Also included are protecting (M1) a pesticidal
CC   polypeptide from proteolytic inactivation in a plant (comprising altering
CC   a proteolytic site within the pesticidal polypeptide that is sensitive to
CC   a plant protease to comprise a proteolytic protection site, where the
CC   proteolytic protection site is not sensitive to the plant protease and
CC   protects the pesticidal polypeptide from proteolytic inactivation in a
CC   plant), an expression cassette (II) comprising the nucleic acid, a
CC   transformed plant (III) a stably incorporated with the expression
CC   cassette in its genome, a transformed seed of the plant and an isolated
CC   pesticidal polypeptide (IV) having proteolytic activity (and comprising
CC   an engineered proteolytic protection site, which is not sensitive to the
CC   plant protease and protects the pesticidal polypeptide from proteolytic
CC   inactivation in a plant). The pesticidal polypeptide is Cry8Bb1 toxin or
CC   its variant or fragment, where the variant and the fragment have
CC   pesticidal activity and the variant has at least 70 % sequence identity
CC   to an amino acid sequence for the Cry8Bb1 toxin. The nucleic acid and/or
CC   expression cassette are useful for protecting a plant from a pest, where
CC   expression of the nucleic acid and/or expression cassette in the plant
CC   produces the pesticidal polypeptide in the plant, and where the
CC   pesticidal polypeptide protects the plant from the pest. The plant
CC   protease is a cysteine protease. The pest is chosen from Colorado potato
CC   beetle, Western corn rootworm, Southern corn rootworm, and boll weevil.
CC   The present sequence encodes a wild-type Bacillus thuringiensis Cry8-like
CC   toxin.
XX
SQ   Sequence 3621 BP; 1277 A; 608 C; 771 G; 965 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:               3.61e-77      Length:         3621
Score:                   948.00        Matches:        248
Percent Similarity:      50.9%         Conservative:   119
Best Local Similarity:   34.4%         Mismatches:     276
Query Match:             26.4%         Indels:         78
DB:                      14            Gaps:           25
```

US-10-782-096-2 (1-682) x AEB45606 (1-3621)

```
Qy           3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
               |||      |||:::|||||||||||||::::::||||||            ::::::::::::
Db           4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy          23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
               ||||||         |||||||||   |||    :::|||          ::::::|||
Db          61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy          43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
               :::||||||::::::|||   |||                      |||:::   |||
Db         121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy          63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                  |||      |||||||||      :::|||   :::||||||         |||
Db         163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy          82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
               :::   ||||||   :::||||||   ||||||||||||
Db         223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy         102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
               :::   ::::::||||||   |||||||||            :::         :::|||
Db         283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy         120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                  |||      |||||||||||||||    :::   |||   :::      |||
Db         343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy         140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
               ||||||   |||||||||||||||         :::||||||   |||   |||:::
Db         403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy         160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                  |||      |||   |||            |||   :::      ::::::
Db         463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy         179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
               |||:::|||:::   |||   :::   ||||||   |||:::      |||||||||:::
Db         523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy         199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                     |||   |||||||||   |||||||||||||||:::|||||||:::||||||
Db         583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642
```

```
Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
Db       2053 AAT 2055
```

RESULT 7
AEB34684
ID   AEB34684 standard; DNA; 3621 BP.
XX
AC   AEB34684;
XX
DT   22-SEP-2005 (first entry)
XX
DE   Bacillus thuringiensis Cry8Bb1 protoxin DNA.
XX
KW   Plant insect pest; toxin; Cry8Bb1 protoxin; insecticide; pesticide; ds.
XX
OS   Bacillus thuringiensis.
XX
PN   WO2005063996-A2.
XX
PD   14-JUL-2005.
XX
PF   20-DEC-2004; 2004WO-US042736.
XX
PR   23-DEC-2003; 2003US-0532185P.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
PA   (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad A, Flannagan RD, Herrmann R, Lu A, Mccutchen BF;
PI   Presnail JK, Rice JA, Wong JF, Yu C;
XX
DR   WPI; 2005-541825/55.
XX
PT   Protection of plants from insect pests comprises the introduction of a
PT   polynucleotide construct that comprises a nucleotide sequence encoding an
PT   insect protoxin; and cleavage of the protoxin by plant protease to form
PT   active insect toxin.
XX
PS   Example 1; SEQ ID NO 5; 65pp; English.
XX
CC   The invention relates to a method of protecting a plant from an insect
CC   pest comprising introducing a polynucleotide construct that comprises a
CC   nucleotide sequence encoding an insect protoxin, operably linked to a
CC   promoter that drives expression in the plant, where the insect protoxin

```
CC   has at least one proteolytic activation site that has been engineered to
CC   comprise a cleavage site that is sensitive to a plant protease (within
CC   the plant), where expression of the construct produces the insect
CC   protoxin in the plant, where cleavage of the protoxin by a plant protease
CC   produces an active insect toxin and where the active insect toxin
CC   protects the plant from the insect pest. The invention also relates to an
CC   isolated nucleic acid molecule comprising a nucleotide sequence encoding
CC   the insect protoxin, an isolated nucleic acid molecule comprising a
CC   nucleotide sequence encoding a Cry8-like insect protoxin, impacting an
CC   insect pest of a plant comprising providing a transgenic plant that
CC   comprises in its genome a stably incorporated polynucleotide construct
CC   and a composition comprising at least one insect protoxin in combination
CC   with a carrier. The composition is useful for impacting insect pests
CC   (e.g. Colorado potato beetle, western corn rootworm, southern corn
CC   rootworm or boll weevil) on plants (e.g. monocot (maize) or dicot). The
CC   modified insect toxins (pro toxin forms of pesticidal polypeptides in
CC   conjunction with endogenous plant or insect gut proteases) are effective
CC   in pest management strategies. This sequence represents Bacillus
CC   thuringiensis Cry8Bb1 protoxin DNA of the invention.
XX
SQ   Sequence 3621 BP; 1277 A; 608 C; 771 G; 965 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.61e-77      Length:        3621
Score:                  948.00        Matches:       248
Percent Similarity:     50.9%         Conservative:  119
Best Local Similarity:  34.4%         Mismatches:    276
Query Match:            26.4%         Indels:        78
DB:                     14            Gaps:          25

US-10-782-096-2 (1-682) x AEB34684 (1-3621)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
```

```
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::   |||      |||:::                           |||   :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         |||||:::||||||||||||||                   ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                 |||   |||            :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::            |||                 |||:::      |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                          :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||      ::::::   ||||||   ||||||
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
              |||
Db       2053 AAT 2055
```

RESULT 8
ABK87247
ID   ABK87247 standard; DNA; 4874 BP.
XX
AC   ABK87247;
XX
DT   07-OCT-2002 (first entry)
XX
DE   Bacillus thuringiensis genomic Cry1218-1 DNA sequence.
XX
KW   Pesticidal; spraying; dusting; broadcasting; seed coating; insect pest;
KW   Colorado potato beetle; western corn rootworm; southern corn rootworm;
KW   insect target range; endotoxin; Cry1218; gene; ds.
XX
OS   Bacillus thuringiensis.
XX
PN   WO200234774-A2.
XX
PD   02-MAY-2002.
XX
PF   24-OCT-2001; 2001WO-US045468.
XX
PR   24-OCT-2000; 2000US-0242838P.
PR   23-OCT-2001; 2001US-00032717.
XX
PA   (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad AR, Duck NB, Feng X, Flannagan RD, Kahn TW, Sims LE;
XX
DR   WPI; 2002-519178/55.
XX
PT   New isolated pesticidal polypeptide useful for impacting insect pest e.g.
PT   Colorado potato beetle.
XX
PS   Claim 13; Page 144-145; 176pp; English.
XX
CC   The present invention relates to a new pesticidal polypeptide. The
CC   invention is useful for impacting an insect pest by applying the the
CC   molecules of the invention to the environment of the insect pest by
CC   spraying, dusting, broadcasting, or seed coating, where the insect pest
CC   is selected from Colorado potato beetle, western corn rootworm or
CC   southern corn rootworm. The invention is also useful for increasing
CC   insect target range and for producing transgenic microorganisms and
CC   plants that express the pesticidal polypeptide. The invention is also
CC   useful for producing transformed plants and in transforming any organism
CC   to produce the pesticidal polypeptide of the invention. The present
CC   nucleic acid sequence encodes a Bacillus thuringiensis wild-type Cry1218
CC   endotoxin protein
XX
SQ   Sequence 4874 BP; 1707 A; 787 C; 1003 G; 1377 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:              5.46e-77      Length:        4874
Score:                  948.00        Matches:       248
Percent Similarity:     50.9%         Conservative:  119
Best Local Similarity:  34.4%         Mismatches:    276
Query Match:            26.4%         Indels:        78
DB:                     6             Gaps:          25

US-10-782-096-2 (1-682) x ABK87247 (1-4874)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||         |||:::||||||||||||::::::||||||         ::::::::::::
Db        734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||         |||||||||   |||   :::|||         ::::::|||
Db        791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                        |||:::   |||
Db        851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||      :::|||   :::||||||            |||
Db        893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||   :::||||||   ||||||||||||
Db        953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||            :::         :::|||
Db       1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||      :::   |||      |||
Db       1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   |||||||||||||||            :::||||||   |||   |||:::
Db       1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||               |||   :::      ::::::
Db       1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::   ||||||   |||:::      |||||||||:::
Db       1253 ATCCTGGATAGTTTATTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                    |||   |||||||||   ||||||||||||||||||:::||||||:::||||||
Db       1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||                  |||   :::   |||   |||
Db       1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                    |||      ||||||   |||||||||:::   :::|||         ||||||
Db       1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492
```

```
Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db   1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db   1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db   1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db   1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db   1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy    359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db   1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy    377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db   1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy    396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db   1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy    416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db   1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy    431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db   2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy    450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db   2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy    470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db   2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy    483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db   2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy    503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db   2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy    523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db   2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy    543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db   2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy    559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db   2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy    572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db   2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy    592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db   2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy    612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db   2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy    632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db   2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy    645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db   2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy    663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db   2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy    682 Asn 682
Db   2783 AAT 2785
```

RESULT 9
ADL15320
ID ADL15320 standard; DNA; 4874 BP.
XX
AC ADL15320;
XX
DT 06-MAY-2004 (first entry)
XX
DE B thuringiensis strain 1218 Cry1218-1 (Cry8Bb1) gDNA.
XX
KW endotoxin; proteolytic site; insect gut; pesticidal; agriculture;
KW plant pest; Cry8-like delta-endotoxin; Cry1218-1; Cry8Bb1; ds.
XX
OS Bacillus thuringiensis.
XX
PN WO2004003148-A2.
XX
PD 08-JAN-2004.
XX
PF 25-JUN-2003; 2003WO-US020027.
XX
PR 26-JUN-2002; 2002US-0391786P.
PR 04-APR-2003; 2003US-0460787P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Abad AR, Flannagan RD, Herrmann R, Kahn TW, Lu AL, Mccutchen BF;
PI Presnail JK, Wong JFH, Yu C;
XX
DR WPI; 2004-083033/08.
XX
PT New endotoxin comprising a mutation consisting of the alteration of at
PT least one proteolytic site, where stability of the endotoxin in insect
PT gut is increased, useful for providing improved pesticidal activity.
XX
PS Example 4; SEQ ID NO 17; 333pp; English.
XX
CC The invention relates to a novel endotoxin having a mutation consisting
CC of the alteration of at least one proteolytic site, where the stability
CC of the endotoxin in an insect gut is increased relative to an endotoxin
CC lacking the mutation. The endotoxin of the invention may be useful for
CC providing improved pesticidal activity against insect pests whilst the
CC compositions of the invention may be useful in agriculture for
CC controlling pests, including plant pests. The current sequence is that of
CC the a DNA related to Bacillus thuringiensis strain 1218 Cry8-like delta-
CC endotoxin Cry1218 of the invention.
XX
SQ Sequence 4874 BP; 1707 A; 787 C; 1003 G; 1377 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:                 5.46e-77      Length:          4874
Score:                     948.00        Matches:         248
Percent Similarity:        50.9%         Conservative:    119
Best Local Similarity:     34.4%         Mismatches:      276
Query Match:               26.4%         Indels:          78
DB:                        12            Gaps:            25

US-10-782-096-2 (1-682) x ADL15320 (1-4874)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db      734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db      791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db     1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db     1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db     1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db     1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db     1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db     1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db     1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db     1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db     1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db     1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db     1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db     1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db     1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db     1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db     1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db     1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy      416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db     1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db     2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db     2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy      470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db     2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db     2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db     2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db     2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db     2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy      559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db     2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db     2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy      592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db     2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy      612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db     2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy      632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db     2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662
```

```
Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::           |||         :::|||:::   |||||||||||::::::
Db       2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||         ::::::   ||||||   ||||||
Db       2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy        682 Asn 682
              |||
Db       2783 AAT 2785

RESULT 10
AEB45622
ID   AEB45622 standard; DNA; 4874 BP.
XX
AC   AEB45622;
XX
DT   22-SEP-2005  (first entry)
XX
DE   B. thuringiensis genomic DNA for Cry1218-1.
XX
KW   Toxin; ds; gene; transgenic plant; insecticide; pesticide;
KW   plant insect pest; Cry8.
XX
OS   Bacillus thuringiensis; strain 1218-1.
XX
PN   WO2005066349-A2.
XX
PD   21-JUL-2005.
XX
PF   09-DEC-2004; 2004WO-US041530.
XX
PR   24-DEC-2003; 2003US-00746914.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad A,  Dong H,  Herrmann R,  Lu A,  Mccutchen BF,  Rice JA;
PI   Schepers EJ,  Wong JF;
XX
DR   WPI; 2005-533612/54.
DR   P-PSDB; AEB45607.
XX
PT   Novel isolated nucleic acid molecule having nucleotide sequence encoding
PT   pesticidal polypeptide comprising engineered proteolytic protection site
PT   resistant to protease, useful for protecting plant from pest e.g.
PT   Colorado potato beetle.
XX
PS   Example 4; SEQ ID NO 17; 362pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule comprising a
CC   nucleotide sequence encoding a pesticidal polypeptide having an
CC   engineered proteolytic protection site, which is not sensitive to the
CC   plant protease and protects the pesticidal polypeptide from proteolytic
CC   inactivation in a plant. Also included are protecting (M1) a pesticidal
CC   polypeptide from proteolytic inactivation in a plant (comprising altering
CC   a proteolytic site within the pesticidal polypeptide that is sensitive to
CC   a plant protease to comprise a proteolytic protection site, where the
CC   proteolytic protection site is not sensitive to the plant protease and
CC   protects the pesticidal polypeptide from proteolytic inactivation in a
CC   plant), an expression cassette (II) comprising the nucleic acid, a
CC   transformed plant (III) a stably incorporated with the expression
CC   cassette in its genome, a transformed seed of the plant and an isolated
CC   pesticidal polypeptide (IV) having proteolytic activity (and comprising
CC   an engineered proteolytic protection site, which is not sensitive to the
CC   plant protease and protects the pesticidal polypeptide from proteolytic
CC   inactivation in a plant). The pesticidal polypeptide is Cry8Bb1 toxin or
CC   its variant or fragment, where the variant and the fragment have
CC   pesticidal activity and the variant has at least 70 % sequence identity
CC   to an amino acid sequence for the Cry8Bb1 toxin. The nucleic acid and/or
CC   expression cassette are useful for protecting a plant from a pest, where
CC   expression of the nucleic acid and/or expression cassette in the plant
CC   produces the pesticidal polypeptide in the plant, and where the
CC   pesticidal polypeptide protects the plant from the pest. The plant
CC   protease is a cysteine protease. The pest is chosen from Colorado potato
CC   beetle, Western corn rootworm, Southern corn rootworm, and boll weevil.
CC   The present sequence is the genomic DNA for a wild-type Bacillus
CC   thuringiensis Cry8-like toxin.
XX
SQ   Sequence 4874 BP; 1707 A; 787 C; 1003 G; 1377 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                5.46e-77      Length:        4874
Score:                    948.00        Matches:       248
Percent Similarity:       50.9%         Conservative:  119
Best Local Similarity:    34.4%         Mismatches:    276
Query Match:              26.4%         Indels:        78
DB:                       14            Gaps:          25

US-10-782-096-2 (1-682) x AEB45622 (1-4874)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||      |||:::||||||||||||::::::||||||         ::::::::::::
Db        734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||            |||||||||   |||      :::|||         ::::::|||
Db        791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                           |||:::   |||
Db        851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||      :::|||   :::||||||               |||
Db        893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::      ||||||   :::||||||   ||||||||||||
Db        953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||         :::            :::|||
Db       1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGAAATT 1072

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      ||||||||||||||      :::   |||   :::      |||
Db       1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||||||||||                  :::||||||   |||   |||:::
Db       1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||                  |||      :::      ::::::
Db       1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::      |||      :::      ||||||   |||:::      |||||||||:::
Db       1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                 |||      |||||||||   ||||||||||||||||::::||||||:::||||||
Db       1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||                  |||   :::   |||   |||
Db       1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                    |||         ||||||   |||||||||:::      :::|||         ||||||
Db       1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                 :::         ::::::|||:::   ||||||   :::|||:::||||||   :::
```

```
Db       1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db       1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db       1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db       1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db       1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db       2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db       2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db       2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db       2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db       2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db       2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy        682 Asn 682
Db       2783 AAT 2785
```

RESULT 11
AAT43221
ID   AAT43221 standard; DNA; 3507 BP.
XX
AC   AAT43221;
XX
DT   16-OCT-2003 (revised)
DT   28-JAN-1997 (first entry)
XX
DE   Antiscarab pest toxin 50C(b) coding sequence.
XX
KW   Toxin 50C(b); scarab pest; toxin; insect; scarabaeidae; pest control;
KW   larval stage insect; grain; tuberous crop; white grub; chafer grub;
KW   cyclocephala; popillia; ds.
XX
OS   Bacillus thuringiensis; strain kumamotoensis.
XX
PN   US5554534-A.
XX
PD   10-SEP-1996.
XX
PF   30-SEP-1994;   94US-00315468.
XX
PR   16-DEC-1991;   91US-00808316.
PR   30-JAN-1992;   92US-00828430.
PR   01-FEB-1993;   93US-00014941.
XX
PA   (MYCO) MYCOGEN CORP.
XX
PI   Foncerrada L, Narva KE, Michaels TE;
XX
DR   WPI; 1996-424659/42.
DR   P-PSDB; AAW06417.
XX
PT   New nucleic acid encoding B.thuringiensis toxin active against scarab(s)
PT   - also related toxin and transformed microbes, effective against adult
PT   pests and their larvae.
XX
PS   Claim 2; Col 27-30; 24pp; English.
XX
CC   AAT43221-T43223 represent the coding sequences for toxins that are active
CC   against scarab pests. This sequence was isolated from the Bacillus
CC   thuringiensis strain kumamotoensis. Insects in the family Scarabaeidae
CC   constitute a serious pest control problem, especially when destructive
CC   larval stage insects infest high value turf found in golf courses,
CC   playing fields and lawns. The larvae of many species also attack grains,
CC   tuberous crops, and ornamentals. The larvae are known as white grubs, or
CC   chafer grubs, and can be found in decaying organic matter, or in the soil
CC   where they consume plant roots. In Europe and the U.S. populations of
CC   these larvae and adults have developed resistance to chemical
CC   insecticides such as the organochlorines and DDT. The toxins encoded by
CC   these sequences, and intact cells that are capable of expressing the

```
CC   proteins, can be used to control many pests of the family scarabaeidae,
CC   such as species of Cyclocephala, and Popillia. The toxins are active
CC   against larvae (present in soil) and against adults. (Updated on 16-OCT-
CC   2003 to standardise OS field)
XX
SQ   Sequence 3507 BP; 1256 A; 576 C; 732 G; 943 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.12e-76      Length:         3507
Score:                  942.50        Matches:        249
Percent Similarity:     50.3%         Conservative:   115
Best Local Similarity:  34.4%         Mismatches:     272
Query Match:            26.2%         Indels:         87
DB:                     2             Gaps:           27

US-10-782-096-2 (1-682) x AAT43221 (1-3507)

Qy         3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db         4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy        23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db        61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy        43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db       121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGTTAGT 162

Qy        63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db       163 GAATACCCTGGTTCACCTGAGGTATTTCTAAGCGAGCAAGATGCAGTTAAGGCCGCAATT 222

Qy        82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db       223 GATATAGTAGGTAAATTACTAACAGGTTTAGGGGTTCCATTTGTTGGGCCGATAGTGAGT 282

Qy       102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGly------ProAspProPheGluAla 119
Db       283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAAAACAAAAGAGTCAATGGGAAATT 342

Qy       120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db       343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy       140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db       403 GCGCTTTCGGAATTGGAAGGGCTAGGGAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy       160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db       463 GAGTGGAAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTTCGAAATCGATTTGAA 522

Qy       179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db       523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTCGAGTGACAAATTTTGAAGTA 582

Qy       199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db       583 CCATTCCTTACAGTATATACAATGGCAGCAAACCTACATTTACTTTTATTAAGGGACGCA 642

Qy       219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db       643 TCAATTTTTGGAGAAGAATGGGGATTGTCTACAAGCACTATTAATAACTACTATAATCGT 702

Qy       239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db       703 CAAATGAAACTTACTGCAGAATATTCTGACCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy       259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db       763 GCAAAATTAAAAGGCTCGAGCGCTAAACAATGGATTGACTATAACCAATTCCGTAGAGAA 822

Qy       279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db       823 ATGACATTGACGGTGTTAGACGTTGTTGCATTATTTTCAAACTATGATACGCGTACGTAT 882

Qy       299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db       883 CCACTGGCAACAACAGCTCAGCTTACAAGGGAAGTATATACAGATCCACTTGGCGCGGTA 942

Qy       319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db       943 GATGTGCCTAATATTGGCTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy       339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db       988 TCAGAAATAGAAAAAGCGGCTATTCGTCCACCTCATGTGTTTGATTATATAACGGGACTC 1047

Qy       359 PheValTyrThrArgThrLeuGlyAsnMetSerAsp---ValArgAsnIleTrpGlyGly 377
Db      1048 ACAGTTTATACAAAAAAACGTAGCTTCACTTCTGATCGTTATATGAGATATTGGGCTGGT 1107

Qy       378 HisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLys 396
Db      1108 CATCAAATAAGCTATAAGCATATCGGTACGAGTAGTACCTTTACACAGATGTATGGAACC 1167

Qy       397 ThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIle 416
Db      1168 AATCAAAATTTACAAAGTACTAGCAATTTTGATTTTACGAATTACGATATTTACAAGACT 1227

Qy       417 GluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsn 431
Db      1228 TTATCAAATGGTGCAGTACTCCTTGATATAGTTTACCCTGGT------------------ 1269

Qy       432 TyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIleAsn 450
Db      1270 TATACGTATACATTTTTTGGAATGCCAGAAACCGAGTTTTTTATGGTAAATCAATTGAAT 1329

Qy       451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
Db      1330 AATACCAGAAAGACGTTAACGTATAAA---CCAGCT---------TCCAAAGATATTATA 1377

Qy       471 Ser--------------------GluLeuProGlyLysAspLysProArgProAsnAla 483
Db      1378 GATCGGACAAGAGATTCGGAATTAGAATTGCCTCCAGAAACTTCAGGTCAACCAAATTAC 1437

Qy       484 GlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSer 503
Db      1438 GAGTCATATAGCCATAGATTAGGTCATATTACATTTATT------------TACTCCAGT 1485

Qy       504 GlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsn 523
Db      1486 TCAACTAGCACGTATGTACCTGTATTTTCTTGGACACATCGGAGTGCAGATCTAACAAAT 1545

Qy       524 ArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIle 543
Db      1546 ACAGTTAAAAGTGGCGAAATCACCCAAATACCAGGG------GGCAAGTCTAGCACCATA 1599

Qy       544 Gly---------PheValIleProGlyPro------ThrGlyGlyAsnLeuValLysVal 558
Db      1600 GGCAGAAATACTTATATAATAAAAGGGCGTGGTTATACAGGGGGAGACTTAGTGGCTTTA 1659

Qy       559 SerAspSerTrpHisSerLeuLysValGlnAla------ProGlnArgGlnThrSerTyr 576
Db      1660 ACGGACCGCATCGGAAGTTGTGAGTTTCAGATGATCTTTCCAGAGTCTCAA---CGATTC 1716

Qy       577 ArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHis 596
Db      1717 CGTATTCGGATTCGTTACGCT--------------------------------------- 1737

Qy       597 SerGlySerSerHisIleValSerPhePheAspCysSerAsnSer--------------- 611
Db      1738 TCTAATGAAACTAGTTATATTAGTTTATACGGACTAAACCAAAGCGGAACTTTAAAATTC 1797

Qy       612 ------------SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIle 627
```

```
Db       1798 AACCAGACATATTCTAATAAAAATGAAAATGATTTAACATATAATGATTTCAAATATATA 1857

Qy        628 AspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArgThrGlnSer 647

Db       1858 GAATATCCAAGAGTCATT------TCAGTAAATGCTTCTTCAAACATACAGAGGTTATCT 1911

Qy        648 PheGlyThrHisAla------------------IleAspLysPheGluPheIleProLeu 661

Db       1912 ATAGGTATACAAACGAATACAAATTTATTTATTTTAGACCGAATCGAATTCATCCCAGTA 1971

Qy        662 Asn---ThrPheProAsnGlnSer---LeuGluLysArgGluGlnGluValAsnAspLeu 679

Db       1972 GATGAGACATATGAAGCGGAAACGGATTTAGAAGCGGCAAAGAAAGCAGTGAATGCCTTG 2031

Qy        680 PheIleAsn 682

Db       2032 TTTACGAAT 2040
```

RESULT 12
ABK51132
ID   ABK51132 standard; cDNA; 3690 BP.
XX
AC   ABK51132;
XX
DT   30-JUL-2002 (first entry)
XX
DE   cDNA encoding Bacillus thuringiensis insecticidal protein.
XX
KW   Insecticide; transgenic; Coleoptera larvae; ss; gene.
XX
OS   Bacillus thuringiensis.
XX
FH   Key   Location/Qualifiers
FT   CDS   187. .3690
FT   /*tag= a
FT   /product= "Insecticide protein"
XX
PN   JP2002045186-A.
XX
PD   12-FEB-2002.
XX
PF   03-AUG-2000; 2000JP-00236140.
XX
PR   03-AUG-2000; 2000JP-00236140.
XX
PA   (SDSB-) SDS BIOTECH CORP.
XX
DR   WPI; 2002-356468/39.
DR   P-PSDB; AAU80281.
XX
PT   A protein having insecticidal activity, a DNA encoding said protein, and
PT   an agent and a method for preventing harmful organisms.
XX
PS   Claim 4; Page 17-18; 19pp; Japanese.
XX
CC   This invention relates to a crystalline protein comprising a fully
CC   defined sequence and the nucleotide sequence encoding this protein. The
CC   protein of the invention is an agent for preventing harmful organisms
CC   comprising Bacillus thuringiensis serovar galleriae SDS502, its mutant or
CC   a microbe transformed by a DNA encoding the protein. This microbe can be
CC   used to produce a protein containing the protein, or containing a protein
CC   having insecticidal activity produced by the SDS502, its mutant or a
CC   transformed microbe, a microbe which is transformed by using the above
CC   DNA and produces the above protein having insecticidal activity, a plant
CC   or a seed transformed by using the above DNA, and Bacillus thuringiensis
CC   serovar galleriae SDS502 producing a protein comprising and producing a
CC   protein showing insecticidal activity.The protein of the invention may
CC   have insecticidal acivity. The agent is used for preventing Coleoptera
CC   larvae. This sequence represents the cDNA encoding the Bacillus
CC   thuringiensis insecticide protein of the invention
XX
SQ   Sequence 3690 BP; 1289 A; 599 C; 762 G; 1040 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:              1.84e-76     Length:         3690
Score:                  940.50       Matches:        256
Percent Similarity:     51.8%        Conservative:   111
Best Local Similarity:  36.1%        Mismatches:     288
Query Match:            26.1%        Indels:         56
DB:                     6            Gaps:           23

US-10-782-096-2 (1-682) x ABK51132 (1-3690)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22

Db        190 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 246

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42

Db        247 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 306

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61

Db        307 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 366

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81

Db        367 GAGACGTTTATTAGT--------------------TCATCTACGGTTCAAACTGGAATT 405

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101

Db        406 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 465

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119

Db        466 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 525

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139

Db        526 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 585

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159

Db        586 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 645

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178

Db        646 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 705

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198

Db        706 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 765

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218

Db        766 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 825

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238

Db        826 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 885

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258

Db        886 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 945

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278

Db        946 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 1005

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298

Db       1006 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 1065

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
```

```
Db       1066 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 1125

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                                |||||||||          |||        ::::::    :::|||
Db       1126 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTTCTTTT 1182

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
              :::   :::|||:::             |||   |||            |||:::    |||
Db       1183 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTTGATATACTCAGTGAAATA 1242

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375
                    ||||||        |||              |||  ::::::           |||
Db       1243 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1302

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
                 ||||||::::::      :::|||   |||      |||   :::      |||:::
Db       1303 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1362

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
              |||   |||   :::            |||:::|||::::::    ||||||      :::|||
Db       1363 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1419

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
                 :::   ||||||                      |||   :::      ||||||
Db       1420 CAGGTTCGATCATTA------------------GGGGCGGATTTAGCTAATTACTACGCA 1461

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
                 |||||||||            |||         :::   |||            |||
Db       1462 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1521

Qy        455 SerLeu---------ArgTyrGluValProAlaAsn---------------LeuProSer 466
              |||:::            |||   |||                                :::
Db       1522 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1581

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
                       :::   |||:::|||   :::::::::|||                    :::
Db       1582 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1635

Qy        487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504
              ||||||||||||||:::|||:::          :::|||            ||||||
Db       1636 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTCTAAGAATGCTAGTAGTCCTGCT 1695

Qy        505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524
                       :::|||      |||   |||   |||   |||   |||   |||
Db       1696 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1755

Qy        525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541
              :::         |||||||||||||||         ||||||                   |||
Db       1756 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1815

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
                                   ||||||   ||||||||||||::::::|||   :::   |||
Db       1816 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1875

Qy        562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577
                          :::   |||:::   |||:::::::|||      |||      ||||||
Db       1876 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1932

Qy        578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597
              |||:::|||||||||            |||         :::||||||      |||
Db       1933 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1986

Qy        598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617
              |||   ||||||::::::      :::   |||   |||   :::               :::
Db       1987 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGGAT 2037

Qy        618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636
                 |||      :::   |||            :::      ||||||      :::
Db       2038 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 2097

Qy        637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653
                 :::      :::               ||||||            |||         :::
Db       2098 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 2157

Qy        654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671
              ||||||||||||   ||||||:::|||   ||||||            :::   |||:::
Db       2158 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2217

Qy        672 ArgGluGlnGluValAsnAspLeuPhe 680
                 ::::::   ||||||   ||||||
Db       2218 GCAAAGAAAGCAGTAAATGGCTTGTTT 2244
```

```
RESULT 13
ABK87236
ID   ABK87236 standard; DNA; 2003 BP.
XX
AC   ABK87236;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Bacillus thuringiensis Cry1218-1 truncated gene sequence.
XX
KW   Pesticidal; spraying; dusting; broadcasting; seed coating; insect pest;
KW   Colorado potato beetle; western corn rootworm; southern corn rootworm;
KW   insect target range; endotoxin; Cry1218; gene; ds.
XX
OS   Bacillus thuringiensis.
XX
PN   WO200234774-A2.
XX
PD   02-MAY-2002.
XX
PF   24-OCT-2001; 2001WO-US045468.
XX
PR   24-OCT-2000; 2000US-0242838P.
PR   23-OCT-2001; 2001US-00032717.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad AR,  Duck NB,  Feng X,  Flannagan RD,  Kahn TW,  Sims LE;
XX
DR   WPI; 2002-519178/55.
DR   P-PSDB; AAU99257.
XX
PT   New isolated pesticidal polypeptide useful for impacting insect pest e.g.
PT   Colorado potato beetle.
XX
PS   Claim 1; Page 106-108; 176pp; English.
XX
CC   The present invention relates to a new pesticidal polypeptide. The
CC   invention is useful for impacting an insect pest by applying the the
CC   molecules of the invention to the environment of the insect pest by
CC   spraying, dusting, broadcasting, or seed coating, where the insect pest
CC   is selected from Colorado potato beetle, western corn rootworm or
CC   southern corn rootworm. The invention is also useful for increasing
CC   insect target range and for producing transgenic microorganisms and
CC   plants that express the pesticidal polypeptide. The invention is also
CC   useful for producing transformed plants and in transforming any organism
CC   to produce the pesticidal polypeptide of the invention. The present
CC   nucleic acid sequence encodes a Bacillus thuringiensis wild-type Cry1218
CC   endotoxin protein
XX
SQ   Sequence 2003 BP; 686 A; 348 C; 392 G; 577 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.73e-75      Length:        2003
Score:                  926.00        Matches:       239
Percent Similarity:     50.8%         Conservative:  116
Best Local Similarity:  34.2%         Mismatches:    268
Query Match:            25.7%         Indels:        76
DB:                     6             Gaps:          23

US-10-782-096-2 (1-682) x ABK87236 (1-2003)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||      |||:::||||||||||||::::::||||||            ::::::::::::
```

```
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||          ||||||||  |||  :::|||           ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||  |||                       |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                |||      |||||||||      :::|||   :::||||||              |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::    ||||||   :::||||||  ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::  ::::::||||||  ||||||||                 :::         :::|||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||     ||||||||||||||||   :::   |||   :::      |||
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||  ||||||||||||||||               :::||||||   |||   |||:::
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||          |||   :::       ::::::
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::   ||||||   |||:::      ||||||||||:::
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                    |||   |||||||||   ||||||||||||||||:::|||||||:::||||||
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||                 |||   :::   |||   |||
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                    |||      ||||||   |||||||||:::      :::|||   ||||||
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                 :::          ::::::|||:::   ||||||   :::|||:::||||||   :::
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   |||||||||:::         ||||||||||||||||||   |||   |||
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||      :::||||||||||||||||||||||||:::|||
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||         |||                ||||||:::          :::|||
Db        943 AACGTGTCTTCAATTGGTTCC-----------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                 :::|||::::::         |||   |||             ::::::      :::
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 ||||||||         |||::::::   :::      :::|||:::   |||
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::             :::   ||||||   :::          :::|||
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                 ::::::                   |||:::|||   |||         ::::::
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                 |||                  |||      :::|||   |||   |||
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||         :::|||:::         |||      |||      :::   :::
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||                |||      |||:::   |||
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                    |||||||||||            :::||||||
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                 :::||||||||||||   :::|||::::::         |||      ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                  |||   |||   |||   |||   |||   |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::      ||||||||||||||||||   |||||||||   |||      |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||               ||||||   |||||||||:::|||:::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
              :::   |||   |||:::                               |||   :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::||||||||||||||||                  ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                 |||   |||          :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::         |||      |||         |||:::            |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644
                       :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

RESULT 14
ABK87241

```
ID   ABK87241 standard; DNA; 2010 BP.
XX
AC   ABK87241;
XX
DT   07-OCT-2002  (first entry)
XX
```

```
DE   Bacillus thuringiensis Cry1218-1A truncated gene sequence.
XX
KW   Pesticidal; spraying; dusting; broadcasting; seed coating; insect pest;
KW   Colorado potato beetle; western corn rootworm; southern corn rootworm;
KW   insect target range; endotoxin; Cry1218; gene; ds.
XX
OS   Bacillus thuringiensis.
XX
PN   WO200234774-A2.
XX
PD   02-MAY-2002.
XX
PF   24-OCT-2001; 2001WO-US045468.
XX
PR   24-OCT-2000; 2000US-0242838P.
PR   23-OCT-2001; 2001US-00032717.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad AR,  Duck NB,  Feng X,  Flannagan RD,  Kahn TW,  Sims LE;
XX
DR   WPI; 2002-519178/55.
DR   P-PSDB; AAU99262.
XX
PT   New isolated pesticidal polypeptide useful for impacting insect pest e.g.
PT   Colorado potato beetle.
XX
PS   Claim 1; Page 123-125; 176pp; English.
XX
CC   The present invention relates to a new pesticidal polypeptide. The
CC   invention is useful for impacting an insect pest by applying the the
CC   molecules of the invention to the environment of the insect pest by
CC   spraying, dusting, broadcasting, or seed coating, where the insect pest
CC   is selected from Colorado potato beetle, western corn rootworm or
CC   southern corn rootworm. The invention is also useful for increasing
CC   insect target range and for producing transgenic microorganisms and
CC   plants that express the pesticidal polypeptide. The invention is also
CC   useful for producing transformed plants and in transforming any organism
CC   to produce the pesticidal polypeptide of the invention. The present
CC   nucleic acid sequence encodes a Bacillus thuringiensis wild-type Cry1218
CC   endotoxin protein
XX
SQ   Sequence 2010 BP; 690 A; 348 C; 394 G; 578 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               1.74e-75       Length:          2010
Score:                   926.00         Matches:         239
Percent Similarity:      50.8%          Conservative:    116
Best Local Similarity:   34.2%          Mismatches:      268
Query Match:             25.7%          Indels:          76
DB:                      6              Gaps:            23

US-10-782-096-2 (1-682) x ABK87241 (1-2010)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326
```

```
Qy         450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db        1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy         470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db        1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy         483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db        1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy         503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db        1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy         523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db        1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy         543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db        1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy         559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db        1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy         572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db        1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy         592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db        1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy         612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db        1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy         632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db        1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy         645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db        1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

RESULT 15
ABK87238

```
ID   ABK87238 standard; DNA; 2010 BP.
XX
AC   ABK87238;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Maize optimised B. thuringiensis Cry1218-1 gene sequence.
XX
KW   Pesticidal; spraying; dusting; broadcasting; seed coating; insect pest;
KW   Colorado potato beetle; western corn rootworm; southern corn rootworm;
KW   insect target range; endotoxin; maize; Cry1218; gene; ds.
XX
OS   Bacillus thuringiensis.
OS   Zea mays.
OS   Synthetic.
XX
PN   WO200234774-A2.
XX
PD   02-MAY-2002.
XX
PF   24-OCT-2001; 2001WO-US045468.
XX
PR   24-OCT-2000; 2000US-0242838P.
PR   23-OCT-2001; 2001US-00032717.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Abad AR,  Duck NB,  Feng X,  Flannagan RD,  Kahn TW,  Sims LE;
XX
DR   WPI; 2002-519178/55.
DR   P-PSDB; AAU99259.
XX
PT   New isolated pesticidal polypeptide useful for impacting insect pest e.g.
PT   Colorado potato beetle.
XX
PS   Claim 1; Page 114-116; 176pp; English.
XX
CC   The present invention relates to a new pesticidal polypeptide. The
CC   invention is useful for impacting an insect pest by applying the the
CC   molecules of the invention to the environment of the insect pest by
CC   spraying, dusting, broadcasting, or seed coating, where the insect pest
CC   is selected from Colorado potato beetle, western corn rootworm or
CC   southern corn rootworm. The invention is also useful for increasing
CC   insect target range and for producing transgenic microorganisms and
CC   plants that express the pesticidal polypeptide. The invention is also
CC   useful for producing transformed plants and in transforming any organism
CC   to produce the pesticidal polypeptide of the invention. The present
CC   nucleic acid sequence encodes a Bacillus thuringiensis wild-type Cry1218
CC   endotoxin protein
XX
SQ   Sequence 2010 BP; 405 A; 837 C; 452 G; 316 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.74e-75      Length:        2010
Score:                  926.00        Matches:       239
Percent Similarity:     50.8%         Conservative:  116
Best Local Similarity:  34.2%         Mismatches:    268
Query Match:            25.7%         Indels:        76
DB:                     6             Gaps:          23

US-10-782-096-2 (1-682) x ABK87238 (1-2010)

Qy           3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db           4 TCCCCCAACAACCAGAACGAGTACGAGATCATCGACGCCACCCCC---TCCACCTCCGTG 60

Qy          23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db          61 TCCAACGACTCCAACCGCTACCCCTTCGCCAACGAGCCCACCAACGCCCTCCAGAACATG 120

Qy          43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db         121 GACTACAAGGACTACCTCAAGATGTCC------------------GCCGGCAACGCCTCC 162

Qy          63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db         163 GAGTACCCCGGCTCCCCCGAGGTGCTCGTGTCCGGCCAGGACGCCGCCAAGGCCGCCATC 222

Qy          82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db         223 GACATCGTGGGCAAGCTCCTCTCCGGCCTCGGCGTGCCCTTCGTGGGCCCCATCGTGTCC 282

Qy         102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db         283 CTCTACACCCAGCTCATCGACATCCTCTGGCCCTCCGGCGAGAAGTCCCAGTGGGAAATC 342

Qy         120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db         343 TTCATGGAGCAGGTGGAGGAGCTCATCAACCAGAAGATCGCCGAGTACGCCCGCAACAAG 402

Qy         140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db         403 GCCCTCTCCGAGCTGGAGGGCCTCGGCAACAACTACCAGCTCTACCTCACCGCCCTGGAG 462

Qy         160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
```

```
Db        463 GAGTGGGAGGAGAACCCCAACGGCTCCCGCGCCCTCCGCGACGTGCGCAACCGCTTCGAG 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTCGACTCCCTCTTCACCCAGTACATGCCCTCCTTCCGCGTGACCAACTTCGAGGTG 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCCTTCCTCACCGTGTACGCCATGGCCGCCAACCTCCACCTCCTCTCCTCAAGGACGCC 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCCATCTTCGGCGAGGAGTGGGGCTGGTCCACCACCACCATCAACAACTACTACGACCGC 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAGATGAAGCTCACCGCCGAGTACTCCGACCACTGCGTGAAGTGGTATGAGACCGGCCTC 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCCAAGCTCAAGGGCACCTCCGCCAAGCAGTGGGTGGACTACAACCAGTTCCGCCGCGAG 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACCCTCGCCGTGCTCGACGTGGTGGCCCTCTTCCCCAACTACGACACCCGCACCTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCCATGGAGACCAAGGCCCAGCTCACCCGCGAGGTGTACACCGACCCGCTCGGCGCCGTG 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCCTCCATCGGCTCT------------TGGTACGAC---AAGGCCCCAAGCTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGCGTGATCGAGTCCTCCGTGATCCGCCCGCCGCACGTGTTCGACTACATCACCGGCCTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACCGTGTACACCCAGTCCCGCTCCATCTCCTCCGCCCGCTACATCCGCCAC---TGGGCC 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGCCACCAGATCTCCTACCACCGCGTGTCCCGCGGCTCCAACCTCCAGCAGATGTACGGC 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACCAACCAGAACCTCCACTCCACCTCCACCTTCGACTTCACCAACTACGACATCTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACCCTCTCCAAGGACGCCGTGCTCCTCGACATCGTGTACCCCGGCTACACC--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TACATCTTC------TTCGGCATGCCGGAGGTGGAGTTCTTCATGGTGAACCAGCTC 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       1327 AACAACACCCGCAAGACCCTCAAATAC------------AACCCCGTGTCCAAGGACATC 1374

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       1375 ATCGCCTCCACCCGCGACTCCGAGCTCGAGCTCCCCCCCGAGACCTCCGACCAGCCCAAC 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       1435 TACGAGTCCTACTCCCACCGCCTCTGCCACATCACCTCCATCCCCGCCACCGGCAACACC 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       1495 ACCGGCCTCGTGCCG---------GTGTTCTCCTGGACCCACCGCTCTGCAGACCTCAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       1546 AACACCATCTACTCCGACAAGATCACCCAGATCCCCGCCGTGAAGTGCTGG---GACAAC 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db       1603 CTCCCCTTCGTGCCCGTGGTGAAGGGCCCCGGCCACACCGGCGGCGACCTCCTCCAGTAC 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db       1663 AACCGCTCCACCGGCTCCGTGGGCACCCTCTTCCTCGCCCGCTACGGCCTCGCCCTGGAG 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db       1723 AAGGCCGGCAAGTACCGCGTGCGCCTCCGCTACGCCACT---------------GACGCC 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       1768 GACATCGTGCTCCACGTGAACGACGCCCAGATCCAG---------------ATGCCCAAG 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       1813 ACCATGAACCCCGGCGAGGACCTCACCTCCAAGACCTTCAAGGTGGCCGACGCCATCACC 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db       1873 ACCCTCAACCTCGCCACCGACTCCTCCCTCGCCCTCAAGCACAACCTCGGCGAGGACCCC 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db       1933 AACTCCACCCTCTCCGGCATCGTGTACGTGGACCGCATCGAGTTCATCCCCGTGGAC 1989
```

Search completed: January 20, 2006, 23:21:12
Job time : 1132 secs

```
                         GenCore version 5.1.7
              Copyright (c) 1993 - 2006  Biocceleration Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:          January 20, 2006, 22:34:18 ; Search time 7058 Seconds
                                          (without alignments)
                                          4520.939 Million cell updates/sec

Title:           US-10-782-096-2
Perfect score:   3597
Sequence:        1 MNSYKNKNEYEMLDALRINS..........TFPNQSLEKREQEVNDLFIN 682

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters:      82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10782096/runat_20012006_095215_24405/app_query.fasta_1.839
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10782096_@CGN_1_1_5315_@runat_20012006_095215_24405 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  gb_est1:*
                 2:  gb_est2:*
                 3:  gb_est3:*
                 4:  gb_htc:*
                 5:  gb_est4:*
                 6:  gb_est5:*
                 7:  gb_est6:*
                 8:  gb_est7:*
                 9:  gb_gss1:*
                 10:  gb_gss2:*
                 11:  gb_gss3:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                      %
Result               Query
  No.      Score     Match Length DB  ID                         Description
-------------------------------------------------------------------------------
    1      121.5      3.4    2006  4  AY850327                   AY850327 Magnaport
    2      117.5      3.3   11493 11  DQ051595                    DQ051595 Homo sapi
c   3        117      3.3     782 10  CL651317                    CL651317 PRI0111d_
    4        113      3.1    2982 10  CL964901                    CL964901 OsIFCC011
    5        110      3.1    1298  1  AJ537963                   AJ537963 AJ537963
    6      109.5      3.0    2823 10  CL967062                    CL967062 OsIFCC014
    7        108      3.0     746 10  CL853210                    CL853210 OR_CBa008
    8        108      3.0    2439 10  CL982610                    CL982610 OsIFSC048
    9      107.5      3.0    1437 10  CL967463                    CL967463 OsIFCC039
c  10      107.5      3.0    1926  9  BH770523                   BH770523 LLMGtag28
   11        107      3.0     500  7  CO516871                   CO516871 s13dSG98H
   12        107      3.0    4979  4  BC032293                   BC032293 Mus muscu
   13        105      2.9     981  2  BI079460                   BI079460 602873864
   14        105      2.9    1016  3  BM904226                   BM904226 AGENCOURT
   15        105      2.9    3180 10  CL961241                    CL961241 OsIFCC005
   16        105      2.9    4379  4  HSM808965                  BX648814 Homo sapi
   17      104.5      2.9    1193  8  DN729834                   DN729834 CNB49-E02
   18      104.5      2.9    2526 10  CL982730                    CL982730 OsIFSC048
   19        103      2.9     576  1  AW776556                   AW776556 EST335621
   20        103      2.9     676  2  BF647300                   BF647300 NF018D07E
   21        103      2.9     692  3  BI308890                   BI308890 EST530300
   22        103      2.9     785  3  BQ165054                   BQ165054 EST610923
   23      102.5      2.8    1152  8  DN656547                   DN656547 CEC23-G01
   24      102.5      2.8    2646 10  CL980167                    CL980167 OsIFCC034
   25        102      2.8    2571 10  AY417032                    AY417032 Mus muscu
   26        102      2.8   72873 11  DQ051072                    DQ051072 Homo sapi
c  27        101      2.8     702  5  BQ874517                   BQ874517 QGI5I09.y
   28      100.5      2.8     830 10  CG459319                    CG459319 PUIKJ96TB
   29      100.5      2.8    1434 10  CL965950                    CL965950 OsIFCC013
   30      100.5      2.8    3072  4  AK084557                   AK084557 Mus muscu
   31      100.5      2.8    3334  4  AK037516                   AK037516 Mus muscu
   32      100.5      2.8    3420  4  AK083596                   AK083596 Mus muscu
   33        100      2.8     607  2  BF651071                   BF651071 NF100E06E
   34        100      2.8     645  6  CB215641                   CB215641 NISC_np08
c  35        100      2.8    1104  2  BG760743                   BG760743 602717027
   36        100      2.8    2094  4  CR626664                   CR626664 full-leng
c  37       99.5      2.8     766  7  CO980294                   CO980294 GM89007B1
   38       99.5      2.8    1000  8  DR152948                   DR152948 47519070
   39       99.5      2.8    1128  8  DN658020                   DN658020 CEC32-C02
   40       99.5      2.8    1266  8  DN702656                   DN702656 CLJ42-H08
   41       99.5      2.8    1626 10  AY419685                    AY419685 Homo sapi
   42       99.5      2.8    1740 10  CL946692                    CL946692 OsIFSB000
   43       99.5      2.8    2109 10  CL967870                    CL967870 OsIFCC016
   44         99      2.8     481  7  CO140775                   CO140775 EST835446
   45         99      2.8     518  7  CO144784                   CO144784 EST819837


                                  ALIGNMENTS

RESULT 1
AY850327
LOCUS       AY850327                2006 bp    mRNA    linear   HTC 01-FEB-2005
DEFINITION  Magnaporthe grisea cytochrome P450-like protein mRNA, complete cds.
ACCESSION   AY850327
VERSION     AY850327.1  GI:58257422
KEYWORDS    HTC.
SOURCE      Magnaporthe grisea (anamorph: Pyricularia grisea)
  ORGANISM  Magnaporthe grisea
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE   1  (bases 1 to 2006)
  AUTHORS   Chen,B.S., Li,Y.Z., Peng,Y.L., Dong,H.T. and Li,D.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-2004) Guangxi University, Daxue Road 100,
            Nanning, Guangxi 530004, China
FEATURES             Location/Qualifiers
     source          1..2006
                     /organism="Magnaporthe grisea"
                     /mol_type="mRNA"
                     /strain="Y34"
                     /db_xref="taxon:148305"
                     /tissue_type="conidia; germ tubes; early appressoria;
                     middle appressoria; late appressoria; mycelia"
     CDS             49..1566
                     /codon_start=1
                     /product="cytochrome P450-like protein"
                     /protein_id="AAW69333.1"
                     /db_xref="GI:58257423"
                     /translation="MSTTQLADVVADRLPLIASAAVILGVTFVVQYFFFKDPLANVPF
                     VGTEFGGEEARREQFLKDAKPFYLEGFRMNKVHKVTSPRTDTTVSIPAEFLEELKRLP
```

```
                 DDTISFGGAVNEHLVKSHLTPALAKINPVLTEEVARTVREELGSLEDWTEVNIHSKLL
                 RIVAIVSGRIFLGPELCHNEKYIESAIYYTMELNQGRQAIFRLPAWQRPFRAWFLPEI
                 KNLHKREREFYALIEPIVKARREARRSDPNYQAPDDMLTWLEESQDKTTRKSIREIAK
                 LQLGLSFAAIHTTTMTTTNIIYDLAAHPEMIAELRQEASAALAAHDGAFTAPALHTMK
                 KMDSVMKESMRMNPAGYSAFMRKVLKPFTLSNGQEIPAGVTIEVPGYGVSRDPETFPN
                 ADSWDGLRFFRLRESGETVGSSKQGGGRASAVEVSAQNQFVSVSKSSLGFGYGRHACP
                 GRFFASNEVKMILARLVLNYDVKLAGGATERYKNIEAAGSCVPDPTKMLLFKERA"
ORIGIN

Alignment Scores:
Pred. No.:              0.0442        Length:          2006
Score:                  121.50        Matches:         135
Percent Similarity:     35.2%         Conservative:    88
Best Local Similarity:  21.3%         Mismatches:      235
Query Match:            3.4%          Indels:          179
DB:                     4             Gaps:            29

US-10-782-096-2 (1-682) x AY850327 (1-2006)

Qy      13 LeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyrProArgTyrProLeuAla 32
Db     250 CTCGAGGGGTTTCGCATGAACAAGGTGCACAAGGTGACGTCGCCGAGG------------ 297

Qy      33 LysAspProGlnMetThrMetArgAsnThrAsnTyrLysGluTrpLeuAsnMetCysAsp 52
Db     298 ACCGACACTACCGTTTCAATCCCCGCCGAGTTCCTGGAGGAGCTCAAGAGGCTGCCCGAT 357

Qy      53 SerAsnThrGlnPheIleGlyAspIleSerThrTyr---------SerSerProGluAla 69
Db     358 GACACCATCAGCTTTGGCGGTGCCGTGAACGAGCACTTGGTCAAGAGCCACCTCACCCCG 417

Qy      70 AlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThrIleLeuSer 89
Db     418 GCCCTCGCCAAAATCAACCCCGTCCTCACCGAA---GAAGTCGCCCGCACCGTCCGCGAG 474

Qy      90 AsnLeuGly---------------------ValProLeuAlaSerGlnSerPheGlyIle 102
Db     475 GAGCTCGGCTCGCTCGAGGACTGGACCGAGGTCAACATCCACAGCAAGCTGCTGCGCATC 534

Qy     103 IleSerArgLeuIleGlyIleLeuTrpAlaGlyProAspProPheGluAlaLeuMetVal 122
Db     535 GTGGCCATCGTGTCGGGGCGCATCTTTCTGGGGCCGGAG---------------CTGTGC 579

Qy     123 LeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArg 142
Db     580 CACAACGAAAAGTACATCGAGTCGGCCATCTACTACACGATGGAGCTCAAC--------- 630

Qy     143 GluLeuGluGlyLeuGlnGlyIleMetArgLeu------TyrGlnThrArgLeuGlnAla 160
Db     631 ------CAGGGCCGCCAGGCCATCTTTCGGCTCCCGGCGTGGCAGCGGCCGTTCCGCGCC 684

Qy     161 TrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThrGlnTyrAlaIleVal 180
Db     685 TGGTTCCTGCCCGAG--------------------------------------------- 699

Qy     181 AspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeu 200
Db     700 ------------ATCAAGAACCTGCACAAG---CGCGAGCGCGAGTTCTACGCCCTCATC 744

Qy     201 LeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyr 220
Db     745 GAGCCCATCGTCAAGGCCCGC------------------------CGCGAGGCCCGGAGG 780

Qy     221 PheGlyAlaGlnTrpGlnLeuGlyAspAsp------------------GluIleArgAsp 234
Db     781 TCCGACCCCAACTACCAGGCCCCCGACGACATGCTCACCTGGCTCGAGGAATCGCAGGAC 840

Qy     235 AsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyr 254
Db     841 AAGACCACGCGCAAGAGC------ATCCGCGAGATT------------------------ 870

Qy     255 AsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArg 274
Db     871 --------------------------GCCAAGCTGCAGCTGGGGCTCAGCTTTGCCGCC 903

Qy     275 PheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAsp 294
Db     904 ATCCACACC---ACCACCATGACCACGACCAACATCATC-------------------- 939

Qy     295 ProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspPro 314
Db     940 ---------TACGACCTCGCCGCCCACCCCGAGATGATCGCCGAGCTGCGCCAAGAGGCC 990

Qy     315 ValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsn 334
Db     991 TCGGCCGCGCTGGCCGCG---CACGACGGCGCCTTCACCGCGCCGGCGCTGCACACCATG 1047

Qy     335 AsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSer---TyrThrThr 353
Db    1048 AAGAAGATGGACTCGGTCATGAAGGAGAGCATGCGCATGAACCCGGCCGGCTACTCGGCC 1107

Qy     354 TrpLeuAsnArgIlePhe---ValTyrThrArgThrLeuGlyAsnMetSerAspValArg 372
Db    1108 TTTATGCGCAAGGTGCTCAAGCCCTTTACG--------------CTGTCCAACGGCCAA 1152

Qy     373 AsnIleTrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHis 392
Db    1153 GAGATCCCCGCCGGCGTCACCATCGAGGTCCCCGGCTACGGCGTGTCGCGCGACCCCGAG 1212

Qy     393 AsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSer 412
Db    1213 ACCTTCCCCAACGCCGACAGCTGGGACGGCCTGCGCTTCTTCAGGCTGCGCGAG------ 1266

Qy     413 ValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyr 432
Db    1267 ------TCGGGCGAGACGGTCGGCAGCTCCAAGCAGGGCGGAGGCCGCGCG--------- 1311

Qy     433 IleThrSerGlnTyrGlyValSerArgValIlePheAsnThrSerAsn------IleAsn 450
Db    1312 --------------------TCCGCCGTCGAGGTCAGCGCCCAGAACCAGTTCGTCAGC 1350

Qy     451 AsnValProGlySerLeuArgTyrGluVal---ProAlaAsnLeuProSerGlnThrIle 469
Db    1351 GTGTCCAAGAGCAGCCTCGGCTTCGG-GTACGGCCGCCACGCCTGCCCCGGCAG---GTT 1406

Qy     470 LeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArg 489
Db    1407 CTTTGCCTCCAACGAGGTCAAGATGATCCTGG-CCA----------------------- 1441

Qy     490 LeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeu 509
Db    1441 ------------------------------------------------------------ 1441

Qy     510 LeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLys 529
Db    1442 ---------GGCTGGTCCTCAACTACGATGTG-AAGCTGGCTGGCGGTGCCACGGAGAGG 1491

Qy     530 IleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPheValIleProGly 549
Db    1492 TACAAGAACATTGAGGCTGCC--------------------GGATCGTGTGTTCCCGAC 1530

Qy     550 ProThrGlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAla 569
Db    1531 CCTACCAAGATGCTTCTT--------------------------TTCAAGGAGCGTGCT 1563

Qy     570 ProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAla---CysLeuValThrHis 588
Db    1564 TGAACTCTGCAGCAGACGACACGATGAAGAATAAATTTCGCGAAACATCTTTGTCGGCAT 1623

Qy     589 GlyAspAlaIlePheValGlu---------------------------------HisSer 597
Db    1624 GGCATTGCGGTGTTTATAGAAGCAGAGCGCGGTTGATTGGGTACGAAGCAGACTTACAAA 1683

Qy     598 GlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db    1684 GGATGTAGCATGCTTGTTTTCTTTTTCTTTTGTGCTTATAGC 1725
```

```
RESULT 2
DQ051595
LOCUS       DQ051595              11493 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Homo sapiens TRRAP gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ051595
VERSION     DQ051595.1  GI:66904805
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 11493)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 11493)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1..11493
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
     gene            <1..>11493
                     /gene="TRRAP"
                     /locus_tag="HC19616"
ORIGIN

Alignment Scores:
Pred. No.:             2.29          Length:         11493
Score:                 117.50        Matches:        157
Percent Similarity:    34.4%         Conservative:   111
Best Local Similarity: 20.2%         Mismatches:     314
Query Match:           3.3%          Indels:         197
DB:                    11            Gaps:           38

US-10-782-096-2 (1-682) x DQ051595 (1-11493)

Qy          2 AsnSerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsn 21
Db       8815 AACCTGGGAAGGAACAACAGCCTGCACGACATGAAGACGGTGGTGAAGACCTGGAGGAAC 8874

Qy         22 MetSerAsnCysTyrProArgTyrProLeuAlaLysAspProGln--------------- 36
Db       8875 ------------------CGACTGCCCATCGTGTCTGACGACTTGTCCCACTGGAGCAGC 8916

Qy         37 ---MetThrMetArgAsnThrAsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThr 55
Db       8917 ATCTTCATGTGGAGGCAGCATCATTACCAGGCGATTGTAACTGCCTATGAGAATAGCTCT 8976

Qy         56 GlnPheIleGlyAspIleSerThrTyrSerSerProGluAlaAlaLeuSerValArgAsp 75
Db       8977 CAG---------------CATGATCCCAGTTCAAATAACGCTATGCTTGGGGTTCAT--- 9018

Qy         76 AlaValLeuThrGlyIleAsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeu 95
Db       9019 GCATCAGCTTCAGCGATCATCCAGTATGGAAAAATCGCCCGGAAACAAGGA------CTG 9072

Qy         96 AlaSerGlnSerPheGlyIleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp 115
Db       9073 GTCAATGTAGCTCTGGATATATTAAGTCGGATTCATACTATT------------------ 9114

Qy        116 ProPheGluAlaLeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArg 135
Db       9115 ------------CCAACTGTTCCTATCGTGGATTGCTTCCAGAAG---ATTCGACAGCAA 9159

Qy        136 ValArg---------------GluAsnAlaLeuArgGluLeuGluGlyLeuGlnGlyIle 150
Db       9160 GTTAAATGCTACCTCCAGCTGGCAGGCGTCATGGGCAAAAACGAGTGCATGCAGGGCCTT 9219

Qy        151 MetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLysAsnAspAspAsnArg 170
Db       9220 GAAGTTATTGAATCTACAAATTTAAAATACTTC-----------------------ACA 9255

Qy        171 ArgAlaLeuValThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLys 190
Db       9256 AAAGAGATGACAGCCGAATTTTATGCACTGAAGGGAATGTTCTTGGCTCAGATCAACAAG 9315

Qy        191 PheLysGluArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeu 210
Db       9316 TCCGAGGAGGCAAAC------------------AAAGCCTTCTCTGCAGCTGTGCAGATG 9357

Qy        211 HisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAsp 230
Db       9358 CACGATGTGCTGGTGAAA---------------------GCCTGGGCCATGTGGGGCGAC 9396

Qy        231 GluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHis--- 249
Db       9397 TACCTGGAGAACATCTTT------------------GTGAAGGAGCGGCAGCTGCACCTG 9438

Qy        250 ---------CysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAla 266
Db       9439 GGCGTGTCTGCCATCACCTGCTACCTGCACGCCTGCCGGCATCAGAACGAGAGCAAATCG 9498

Qy        267 GlnAspTrp---------------------ValSerPheAsnArgPheArgThrAspMet 279
Db       9499 AGGAAATACTTAGCCAAGGTGCTGTGGCTTTTGAGTTTTGATGATGACAAAAACACTTTG 9558

Qy        280 ThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrPro 299
Db       9559 GCAGATGCCGTCGACAAGTACTGCATTGGTGTGCCA---------CCCATCCAGTGGCTG 9609

Qy        300 LeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGly 319
Db       9610 GCCTGGATCCCACAGCTGCTCACCTGCCTGGTTGGCTCGGAGGGAAAGCTGCTCTTGAAC 9669

Qy        320 ValLeuGluSerGlyGlyArgThrTyrProTrp------TyrAsnProAsnAsnThrThr 337
Db       9670 CTCATTAGCCAGGTTGGACGCGTGTATCCCCAAGCGGTCTACTTTCCCATCCGGACCCTG 9729

Qy        338 PheThrAlaMetGluAsnAsnAlaArgArgArg------------------ProSerTyr 351
Db       9730 TACCTGACCCTGAAAATAGAACAGCGGGAACGCTACAAGAGCGATCCAGGGCCCATAAGA 9789

Qy        352 ThrThr---------Trp---LeuAsnArgIlePheValTyrThrArgThrLeu------ 365
Db       9790 GCAACAGCACCCATGTGGCGCTGCAGCCGAATCATGCACATGCAGCGAGAGCTCCACCCC 9849

Qy        366 ------------------------------GlyAsnMetSerAspValArgAsnIleTrp 375
Db       9850 ACCCTTCTGTCTTCCCTGGAAGGCATCGTCGATCAGATGGTCTGGTTCAGAGAAAATTGG 9909

Qy        376 GlyGlyHisThrLeu------ValGluAsnGly--------------------------- 384
Db       9910 CATGAAGAGGTTCTCAGGCAGCTCCAACAGGGCCTGGCGAAATGTTACTCCGTGGCGTTT 9969

Qy        385 ------------------AsnAspGlySerGluIleThrHis------AsnPhe---Gly 395
Db       9970 GAGAAAAGTGGAGCGGTGTCCGATGCTAAAATCACCCCCCACACTCTCAATTTTGTGAAG 10029

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe--- 414
```

```
                   |||       |||          :::          |||    :::|||:::|||
Db      10030 AAGTTGGTGAGCACGTTTGGGGTGGCCTGGAGAATGTGTCCAACGTCTCGACCATGTTC 10089

Qy        415 ---------SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsn 431
                       :::  |||||||||||||
Db      10090 TCCAGCGCAGCCTCTGAGTCTCTGGCCCGG-----------------------CGGGCG 10125

Qy        432 TyrIleThrSerGlnTyrGlyValSerArgValIle---------PheAsnThrSerAsn 448
                  |||:::|||        |||   :::   :::            |||   |||
Db      10126 CAGGCCACTGCACAAGACCCTGTCTTTCAGAAGCTGAAAGGCCAGTTCACGACGGATTTT 10185

Qy        449 IleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSer------ 466
                 :::||||||||||||:::::                  ||||||   |||
Db      10186 GACTTCAGCGTTCCAGGATCCATGAAGCTTCAT---------AATCTTATTTCTAAGTTG 10236

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
              :::      :::      |||      |||   |||      |||
Db      10237 AAAAAGTGGATCAAAATCTTGGAGGCCAAGACCAAGCAACTCCCCAAATTCTTCCTCATA 10296

Qy        487 SerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIle 506
                 :::      ::::::|||||||||   |||:::   |||   ||||||
Db      10297 GAGGAAAAGTGCCGGTTCTTGAGCAATTTCTCGGCACAGAC-AGCTGAAGTGGAAAT--- 10352

Qy        507 ValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGlu 526
                                    |||      :::         |||      |||
Db      10353 ------------------TCCTGGGGAGTTTCTGATGCCAAAGCCAACGCATTATTACAT 10394

Qy        527 ProAspLysIleThrGlnIle---------AspAlaValLysGlyTrpGlyGlyAsnIle 543
                 |||                           |||      :::         :::
Db      10395 CAAGATTGCACGGTTCATGCCCCGGGTAGAGATTGTGCAGAAGCACAACACCGCAGCCCG 10454

Qy        544 GlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSerTrpHis 563
                    |||  ||||||                                        |||
Db      10455 GCGGCTGTACATCCGGGG-----------------------------ACACAATGGCAA 10484

Qy        564 SerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAla 583
                 |||   :::         :::|||:::      :::|||:::      |||   |||
Db      10485 GATCTACCCATACCTCGTCATGAACGACGCCTGCCTCACAGAGTCACGGCGAGAGGAGCG 10544

Qy        584 CysLeuValThrHisGlyAspAla------IlePheValGluHisSerGlySerSerHis 601
              |||:::            |||         :::|||   |||         |||
Db      10545 TGTGTTGCAGCTGCTGCGTCTGCTGAACCCCTGTTTGGAGAAGAGAAAGGAGACCACCAA 10604

Qy        602 IleValSerPhePheAspCysSerAsnSerSerGlyArg---ProSerAsnThr-LeuLe 620
                          |||         :::         |||      |||::::::     |||::
Db      10605 GAGGCACTTGTTTTTCACAGTGCCCCGGGTTGTGGCAGTTTCCCCACAGATGCGCCTCGT 10664

Qy        620 uGluSerAspPheArgTyrIleAspValProGlyIlePheThrProSerIle-------- 637
              :|||   :::         :::   :::      |||:::
Db      10665 GGAGGACAACCCCTCTTCACTTTCCCTTGTGGAGATCTACAAGCAGCGCTGCGCCAAGAA 10724

Qy        638 ----------------AsnProLeuIleArgTyr---------------ArgThrGlnSe 647
                              ||||||:::   ||||||                      |||::
Db      10725 GGCATCGAGCATGACAACCCCATCTCCCGTTACTATGACCGGCTGGCTACGGTGCAGGC 10784

Qy        647 rPheGlyThrHisAlaIleAspLysPhe-----------------------------Gl 657
              :   ||||||   |||         :::
Db      10785 GCGGGGAACCCAAGCCAGCCACCAGGTCCTCCGCGACATCCTCAAGGAGGTTCAGAGTAA 10844

Qy        657 uPheIlePro-----------------------LeuAsnThrPheProAsn 666
                  :::|||                          |||:::||||||||||||
Db      10845 CATGGTGCCGCGCAGCATGCTCAAGGAGTGGGCGCTGCACACCTTCCCCAAT 10896
```

RESULT 3
CL651317/c

```
LOCUS       CL651317               782 bp    DNA     linear   GSS 09-JUL-2004
DEFINITION  PRI0111d_D12 - PRI0111d.B21 (782) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL651317
VERSION     CL651317.1  GI:50129904
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 782)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE     AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
   PUBMED   14681447
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            This library was generated at Caltech, Pasadena, USA and end
            sequenced at Vancouver, Canada.
            Seq primer: T7
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1..782
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN

Alignment Scores:
Pred. No.:              0.0316          Length:          782
Score:                  117.00          Matches:         68
Percent Similarity:     36.8%           Conservative:    37
Best Local Similarity:  23.9%           Mismatches:      104
Query Match:            3.3%            Indels:          76
DB:                     10              Gaps:            14

US-10-782-096-2 (1-682) x CL651317 (1-782)

Qy        364 ThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly---GlyHisThrLeuValGlu 382
              |||||||||||         |||   :::|||            :::||||||::::::
Db        679 ACGTTGGGCAATGCCGGAGGCGTTGTTGATATCGGTACGTATGAATATACCTTGCTGGAT 620

Qy        383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLys---ThrAspSerIleThr 401
              ||||||||            :::      |||      :::   |||   |||   |||
Db        619 AATGGCAACCATAGCTGGAGTCTGGCAGAGAATCGCGCGCAAATTACCCCTTCAACCACT 560

Qy        402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGlu---SerLeuAla 420
                 :::                     |||      ||||||         |||
Db        559 GATGTGCTGAATATGGCGGCCGCACAACCGCTGGTATTTGATGCAGAACTGGACACCGTG 500

Qy        421 ArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSer 440
              |||      ||||||            :::   :::|||   |||:::   :::   :::
Db        499 CGTGAGCGTCTTGGTAGCGTAAAAGGCGTTAGTTACGATACGGCGATGTGGAGTTCGGCA 440

Qy        441 ArgValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluVal 460
                             ||||||   |||                        :::|||
Db        439 ---------ATTAACACCCGCAACAACGTGACCACTGATGCGGGAGCTGGTTTTGAG--- 392

Qy        461 ProAlaAsnLeuProSerGlnThrIleLeuSerGluLeuProGlyLysAspLysProArg 480
                                 ||||||:::                  |||   |||
Db        391 -----------------CAAACATTGACGGGCCTGACGCTCGGTATAGAT--------- 359

Qy        481 ProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArg 500
                                                          |||   |||
Db        358 ------------------------------------AGCCGTTTCTCCCGTGAAGAA 338

Qy        501 SerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMet--- 519
              ||||||:::            :::      |||   |||||||::::::|||:::   :::
Db        337 AGCAGTACA---------ATTCGCGGCTTGATCTTTGGTTACTCTCATTCTGATATTGGT 287
```

```
Qy        520 ---AspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGly 538
                 ||||||   :::         |||     :::    |||       |||
Db        286 TTTGATCGCGGCGGCAAAGGTAATATCGATAGCTATACC---CTGGGGGCTTATGCCGGT 230

Qy        539 TrpGlyGlyAsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysVal 558
              |||         |||     :::   |||                 :::|||||||||
Db        229 TGGGAGCATCAGAACGGTGCCTATGTTGATGGG---------------GTGGTGAAAGTT 185

Qy        559 SerAspSerTrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIle 578

Db        184 GAC--------------------------------------------------------- 182

Qy        579 ArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGly 598
                      |||:::|||        :::                          |||
Db        181 ------CGTTTTGCCAACACCATC--------------------------CATGGCAAG 155

Qy        599 SerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThr 618
                 |||:::       :::    |||        ||||||:::|||       ::::::
Db        154 ATGAGTAATGGGGCAACAGCGTTTGGCGATTACAATAGTAACGGCGCGGGTGCTCAT--- 98

Qy        619 LeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsn 638
                 :::||||||   ||||||:::::|||       |||            |||:::
Db         97 ---GTTGAGAGCGGGTTCCGTTGGGTTGACGGATAGGGG-----------AGTGTTAGA 53

Qy        639 ProLeuIleArgTyr 643
              |||   :::   :::
Db         52 CCCTATCTGGCCTTT 38
```

RESULT 4
CL964901

```
LOCUS       CL964901               2982 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFCC011268 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL964901
VERSION     CL964901.1  GI:52384490
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 2982)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1..2982
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Alignment Scores:
Pred. No.:              0.839          Length:          2982
Score:                  113.00         Matches:         137
Percent Similarity:     32.2%          Conservative:    73
Best Local Similarity:  21.0%          Mismatches:      204
Query Match:            3.1%           Indels:          238
DB:                     10             Gaps:            32
```

US-10-782-096-2 (1-682) x CL964901 (1-2982)

```
Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              |||         :::|||||||:::         |||                |||:::|||
Db        376 AACCAACTGAGCTACCTCAACCTC------AGCACGAACTCTTTGGAAGGTAACATCCCA 429

Qy         63 ThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsn 82
              :::         |||     :::   |||   :::   |||         |||:::      |||
Db        430 TCTGAGCTCTCTTCATGTTCCCAGCTTAAAATTCTTGAC------CTTAGCAACAACAAT 483

Qy         83 SerValGlyThrIleLeuSerAsnLeuGly---ValProLeuAlaSerGlnSerPheGly 101
                    |||:::|||   |||      |||      :::||||||         :::
Db        484 CTCCAAGGGAGCATCCCTTCTGCTTTTGGGGACTTACCTTTGCTACAAAAACTAGTCCTT 543

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAspProPhe---GluAlaLeu 120
                    |||||||||   |||      :::                |||:::         ::::::
Db        544 GCCAACAGCAGGCTTGCAGGTGAAATA---------CCAGAGTCTTTAGGAAGCAGCATA 594

Qy        121 MetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAla 140
                 :::         :::|||         ::::::         |||:::      |||:::
Db        595 TCGCTCACATATGTTGACCTTGGAAACAACGCTCTCACAGGGAGAATCCCAGAGTCCTTG 654

Qy        141 LeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAla 160
              :::                  ||||||   :::::|||||||      :::      |||
Db        655 GTAAATAGT---TCATCTCTACAA---GTGCTTAGGCTCATGAGAAATGCCCTTAGCGGA 708

Qy        161 TrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThrGlnTyrAlaIleVal 180
                 |||         |||      |||      :::|||   |||            :::
Db        709 CAACTCCCAACT---AATATGTTCAACTCGTCGTCTCTT---ACTGACATTTGCCTCCAA 762

Qy        181 AspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeu 200
                 |||   |||         :::|||
Db        763 CAGAACAGCTTTGGTGGTACTATACCA--------------------------------- 789

Qy        201 LeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyr 220
                 ||||||   |||   ::::::   :::      :::   |||   |||   :::
Db        790 ---CCTGTTACTGCCATGTCTTCACAAGTCAAGTATCTTGATTTATCTGACAACAATCTC 846

Qy        221 PheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGln 240
                 |||                     :::                  |||:::||||||
Db        847 ATTGGAACAATGCCTTCCTCAATAGGAAACCTTTCTTCGCTAATCTATGTTCGTCTTTCA 906

Qy        241 ---------------GlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsn 255
                             |||   |||   |||         |||         |||
Db        907 CGAAACATTTTACTCGGGAGCATTCCAGAGAGCTTAGGACATGTCGCAACACTTGAGGTT 966

Qy        256 GlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPhe 275
                       ||||||      |||
Db        967 ATATCCTTGAATAGCAACAATCTG------------------------------------ 990

Qy        276 ArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspPro 295

Db        990 ------------------------------------------------------------ 990

Qy        296 ArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProVal 315

Db        990 ------------------------------------------------------------ 990

Qy        316 GlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsn 335
                                   ||||||                  :::|||      ::::::
Db        991 --------------------TCAGGGTCAGTACCCCAATCTCTCTTCAACATGTCCTCT 1029

Qy        336 ThrThrPheThrAlaMetGluAsnAsnAla------ArgArgArgProSerTyrThrThr 353
                 ||||||   ||||||   ||||||:::               :::   ||||||
Db       1030 CTGACATTCCTTGCCATGACAAATAACTCACTCATTGGAAAAATTCCTTCAAACATCGGC 1089

Qy        354 TrpLeuAsnArgIlePheValTyrThrArgThrLeuGlyAsn---------------Met 368
              :::                                 ||||||   |||               :::
Db       1090 TAC--------------------------ACTCTCCCAAATATTCAGGAATTGTACTTG 1122

Qy        369 SerAspValArgAsnIleTrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySer 388
              |||||||||
```

```
Db     1123 TCAGATGTC--------------------------------------------------- 1131

Qy      389 GluIleThrHisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPhe 408
                                     |||   |||         |||            |||
Db     1132 ---------------------AAGTTTGATGGTTCGATCCCAGCTTCACTTCTCAATGCA 1170

Qy      409 AlaAsnLeuSerValPhe------------------------------------------ 414
            :::||||||      |||
Db     1171 TCAAATCTACAGACTTTTAACCTGGCAAACTGTGGCCTAACTGGGTCCATTCCACTGTTA 1230

Qy      415 ---SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThr------GluAlaAsnAsn 431
               |||:::   :::|||   ::::::   |||||             ||||||:::
Db     1231 GGGTCATTGCCAAATTTGCAGAAACTTGATTTAGGATTCAACATGTTTGAGGCAGATGGA 1290

Qy      432 Tyr---------------IleThrSerGlnTyrGlyValSerArgValIlePheAsnThr 446
            :::               :::|||:::         ::::::|||::::::   :::
Db     1291 TGGAGCTTTGTCTCTTCACTCACTAATTGCTCTAGGTTGACTAGACTGATGCTGGATGGG 1350

Qy      447 SerAsnIle--------------------------------------------------- 449
            :::||||||
Db     1351 AACAATATCCAAGGGAACCTGCCAAGTACTATTGGAAATCTTTCTAGTGATCTACAGTGG 1410

Qy      450 ---------------AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeu 464
                           ||||||:::   ||||||:::      |||:::      ||||||
Db     1411 TTGTGGCTAGGCGGCAACAACATTTCTGGGTCTATACCACCAGAGATT---GGTAATCTC 1467

Qy      465 ProSerGlnThrIleLeu------------SerGluLeuProGlyLysAspLysProArg 480
                     |||   |||            :::   |||   |||         |||
Db     1468 AAGGGCCTCACTAAGTTGTACATGGATTACAATCTTTTGACTGGAAATATACCACCAACA 1527

Qy      481 ------------ProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPhe 496
                           |||      |||   :::         :::|||:::|||
Db     1528 ATTGGAAATTTACACAACTTGGTCGATATAAATTTCACACAAAATTATCTCTCAGGAGTG 1587

Qy      497 ------AspAla---------------------------------------ArgArgSer 501
                  ||||||                                       |||   :::
Db     1588 ATCCCAGATGCTATTGGAAATCTTTTGCAACTGACTAACCTGAGACTAGATAGGAACAAC 1647

Qy      502 SerSerGlyGlyIle---------------------ValSerLeuLeuThrPheGlyTrp 514
               ||||||   |||                           |||   |||
Db     1648 TTTAGTGGAAGTATACCTGCAAGCATAGGCCAGTGCACTCAACTTACAACACTCAACCTT 1707

Qy      515 AlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIle--- 533
            |||:::   |||:::         |||         |||   ||||||   ||||||
Db     1708 GCTTACAACTCATTA---------AATGGGAGCATACCGAGTAAAATCTTTCAAATTTAT 1758

Qy      534 ------------------AspAlaValLysGlyTrp------GlyGlyAsnIleGlyPhe 545
                              |||            :::      ||||||
Db     1759 CCTCTTTCTGTTGTTTTGGATTTGTCACACAATTACTTGTCTGGAGGT------------ 1806

Qy      546 ValIleProGlyProThrGlyGly------AsnLeuValLysValSerAspSerTrpHis 563
               ||||||         |||         ||||||||||:::|||      |||   :::
Db     1807 ---ATTCCAGAGGAAGTTGGCAACCTCGTTAATTTGAATAAACTCAGCATCTCAAACAAT 1863

Qy      564 SerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAla 583
               |||      :::   |||
Db     1864 AGGTTGTCTGGTGAAGTTCCATCC------------------------------------ 1887

Qy      584 CysLeuValThrHisGlyAspAlaIlePheValGlu 595
                     |||   |||:::   :::   :::|||
Db     1888 ---------ACTCTAGGCGAATGTGTGCTTCTAGAG 1914
```

RESULT 5
AJ537963

```
LOCUS       AJ537963                1298 bp    mRNA    linear   EST 01-JUL-2004
DEFINITION  AJ537963 1A1 Timarcha balearica cDNA clone Timarcha5F4, mRNA
            sequence.
ACCESSION   AJ537963
VERSION     AJ537963.1  GI:40310965
KEYWORDS    EST.
SOURCE      Timarcha balearica
  ORGANISM  Timarcha balearica
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
            Chrysomeloidea; Chrysomelidae; Chrysomelinae; Timarchini; Timarcha.
REFERENCE   1  (bases 1 to 1298)
  AUTHORS   Gomez-Zurita,J., Kopliku,F., Theodorides,K. and Vogler,A.P.
  TITLE     Resources for a phylogenomic approach in leaf beetle (Coleoptera)
            systematics
  JOURNAL   (in) Jolivet,P., Santiago-Blay,J.A. and Schmitt,M. (Eds.); NEW
            DEVELOPMENTS IN THE BIOLOGY OF CHRYSOMELIDAE, 19-35 (2004)
COMMENT     Contact: Gomez-Zurita J
            Department of Entomology
            The Natural History Museum
            Cromwell Road, London SW7 5BD, United Kingdom.
FEATURES             Location/Qualifiers
     source          1..1298
                     /organism="Timarcha balearica"
                     /mol_type="mRNA"
                     /db_xref="taxon:79517"
                     /clone="Timarcha5F4"
                     /dev_stage="adult"
                     /clone_lib="1A1"
                     /note="country: Spain:Balearic Islands, Mallorca,
                     Esporles"
ORIGIN

Alignment Scores:
Pred. No.:              0.48        Length:         1298
Score:                  110.00      Matches:        69
Percent Similarity:     39.2%       Conservative:   52
Best Local Similarity:  22.3%       Mismatches:     98
Query Match:            3.1%        Indels:         90
DB:                     1           Gaps:           19

US-10-782-096-2 (1-682) x AJ537963 (1-1298)

Qy      163 ValAsnLysAsnAspAsp------------AsnArgArgAlaLeu--------------- 173
            |||:::||||||::::::            |||   |||
Db      144 GTGAGCAAGAATAATAACGAGTTCAAAGTCAATAACCGGATATTCTTTTTGAATCATAAT 203

Qy      174 ValThrGlnTyrAlaIle------------ValAspAsnPhePheGluLysAsnMetPro 189
            :::|||   |||                  ||||||||||||         |||:::
Db      204 ATCACAGATTACGATTATAATTATAAAAAAGTAGATAATTTCAAGCACGCAAACGTTGTC 263

Qy      190 LysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsn 209
                  :::|||::::::   |||
Db      264 ACAACAGAAGAAAAAAGTAAGGAATAC--------------------------------- 290

Qy      210 LeuHisLeuIleLeuLeuArgAspAlaAspTyrPhe---GlyAlaGlnTrpGlnLeuGly 228
                                 |||:::   ||||||         :::|||
Db      291 ---------------------GATTCGAAATATTTCCTGGACGAAGAATGGAACTACGAT 329

Qy      229 AspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArgGlu---TyrLys 247
               :::   |||      :::   |||   |||         |||   :::   ||||||
Db      330 CACGAAAACATC------TCAAAAATCGAATTGCCACTGTTTATCCACAAATTCTATAAA 383

Qy      248 Asp---------------------------HisCysIleThrPhe---TyrAsnGlnGly 257
            |||                           |||   :::   |||   :::|||
Db      384 GATCAATCCGAAGACAAAACGTCCGATAAGCACACTTTGCGCTTTTTATTCAACACCAAT 443

Qy      258 LeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThr 277
               |||:::   :::|||                  ::::::   |||      :::
Db      444 AAAAACAAATCGAGCAGGATCATAGAT---------TTGAACCTTAACACGTTACAGCCG 494

Qy      278 AspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArg 297
            :::   |||:::         ||||||         :::|||
Db      495 GAAGAAACCATTGTAGACGCCGATTTATACTTCTATTGGCCT------------------ 536

Qy      298 TyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPhe 317
                                 :::   |||      :::|||   :::   |||
Db      537 ------------------GTCCAAAACACCTCGGCAATATATAAAGAATCTGTA------ 572

Qy      318 ThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThr 337
                  ||||||            |||   |||               :::
Db      573 ------GTGCTA------------AGGCTTTACCAAATTGAAAGACAATCCACCGGAGAT 614
```

```
Qy        338 PheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArg 357
                |||         |||::::::::    |||   |||              ::::::::
Db        615 TTTAAT------GAATCGGACTCAATACGAAATCCTGAT-----------GTTCACAAG 656

Qy        358 IlePheValTyrThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGly 377
                :::|||              |||:::    :::|||    |||     |||
Db        657 CTGTTC------------ACCGTTATTTATGTATCTAAGGCGCGAAAGGCGTGGATGAAA 704

Qy        378 HisThrLeu---------ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
                   |||:::          ::::::|||   :::|||   |||        |||:::
Db        705 TTCACAATAAAGAAACCGATCGATAACTGGTTGAACGGGGAAGAA---------AACTGG 755

Qy        395 Gly-----------------LysThrAspSerIleThrProIleGlnTyrPheAsnPhe 408
                |||                          |||:::   |||   |||:::   ||||||
Db        756 GGTTTGCTTCTTACCATATCTAACTACGATAATAATACATTAATAGAAGTTTTTAATGAC 815

Qy        409 AlaAsnLeuSerValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGlu 428
                   |||       ::::::|||      ::::::   ::::::           ||||||
Db        816 ACTAACGATGGCATTTATAGTACGTTTGCTGTAGTAAAAGTTCAAAGAAATGATACAGAA 875

Qy        429 AlaAsnAsnTyrIleThrSerGlnTyr 437
                   |||::::::   ||||||   :::
Db        876 ACAAATTCTTTTGATACTTCAACCCAC 902
```

```
RESULT 6
CL967062
LOCUS       CL967062                2823 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFCC014690 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL967062
VERSION     CL967062.1  GI:52388772
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 2823)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1..2823
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Alignment Scores:
Pred. No.:              1.97            Length:         2823
Score:                  109.50          Matches:        158
Percent Similarity:     30.3%           Conservative:   102
Best Local Similarity:  18.4%           Mismatches:     249
Query Match:            3.0%            Indels:         348
DB:                     10              Gaps:           42

US-10-782-096-2 (1-682) x CL967062 (1-2823)
```

```
Qy         45 LysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyr 64
              |||:::    |||   ||||||||                        |||   |||
Db        480 AAAAAACTTCCTAACTGGTGTGATTCC----------------------TCCTAGTAT 515

Qy         65 SerSerProGluAlaAla-LeuSerValArg------AspAlaValLeuThrGlyIleAs 82
                 |||      |||  :::    :::|||          |||||||||
Db        516 AGACTCCCTCTCCAAGCTTATTGGTTTGAGACTTGGCCAGAACAATCTTACTGGG----- 570

Qy         82 nSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIl 102
               ::::::    |||:::|||:::|||:::                |||:::      ::
Db        571 -ACCATCCCAACAATGCTCACAAACATC---------ACTACTTTACAAACACTCGATCT 620

Qy        102 eIle---------------SerArgLeuIleGlyIleLeuTrpAlaGlyProAspProPh 117
              :|||                  ::::::       ||||||      |||:::
Db        621 CATATACGGTCAGCTCACCGGAAAAATTCCTGACGAGCTTTGGCAAATGCCAAAC----- 675

Qy        117 eGluAlaLeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValAr 137
                                 :::|||||||||
Db        676 ------------------ATAGAGGAATTG------------------------------ 687

Qy        137 gGluAsnAlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrAr 157
               :::                   |||   |||      ||||||   :::   :::
Db        688 -GATCTAGGTGGAAATAACCTATCAGGT------GGGATTTCACAAACTTTT-------- 732

Qy        157 gLeuGlnAlaTrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThrGlnTy 177
                                   |||      :::   :::   |||
Db        733 ---------------------CCTAATATATCTTCCCTCCAATATCTAGACTTGGCGAT 770

Qy        177 rAlaIleValAspAsnPhePheGluLysAsn-----------MetPro----------- 189
                 ::::::||||||                  |||         :::|||
Db        771 AAATATGCTAGACAACACATTGCCGTCTAACATTGGCGATGCTCTCCCCAACCTCGTATA 830

Qy        190 ----LysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAl 208
                           |||   :::||||||       :::                  |||::
Db        831 TATATCCTTAGCAGAAAACTCCTTCGAAGGTCAAATCCCATCTTCCCTAGGAAATGCTTC 890

Qy        208 aAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGl 228
              :   |||      |||   |||         ::::::   |||
Db        891 AGGATTATCTCAAATAGACCTGTCCTCTAATAATTTCAATGGT----------------- 933

Qy        228 yAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAs 248
                     ::::::      :::            |||   ||||||||||||
Db        934 -------CAAGTTCCAAGCTCTTTGGGGATGCTTTCAGGGCTGATTCGC----------- 975

Qy        248 pHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAs 268
                                  |||      |||      |||   |||:::   :::
Db        976 ---------------CTAAACGTTGAGAGCAATATGCTTAATGCAAGTGATAATGAAGG 1019

Qy        268 pTrpValSerPheAsn--------------ArgPheArgThrAspMetThrLeuThrVa 283
               |||      ||||||              |||            ::::::   |||
Db       1020 CTGGGATTTCTTTAATGCGTTAAGCAACTGCCGTAATCTAGAGCAGCTTTCATTCACGGA 1079

Qy        283 lLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLy 303
                       |||         :::   ||||||
Db       1080 TAATCAGCTCCAAGGAGTCATACCAAAT-------------------------------- 1107

Qy        303 sThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSe 323
                                                   ||||||      |||
Db       1108 ---------------------------------TCAGTTGGCAACCTGTCTGTC-------- 1128

Qy        323 rGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAs 343
                                                  |||      :::   |||
Db       1129 --------------------------------AATCTTCTATATCTACTAATGAGTGA 1154

Qy        343 nAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrAr 363
               |||                                             |||
Db       1155 AAATAAATTATCAGGT--------------------------------GTAGTTCCCCC 1181

Qy        363 gThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHisThrLeuValGluAs 383
               ::::::||||||::::::                           :::|||:::|||
Db       1182 AAGCATAGGAAACCTTACA----------------------------AGCTTAATTGAACT 1214

Qy        383 nGlyAsnAspGlySerGluIleThr------------HisAsnPheGlyLysThrAspSe 399
```

```
                |||   |||   :::   :::|||                          ||||||   ::::::
Db       1215 CGGGCTAGATGCAAACAGCCTTACTGGTCCAATTGATGAGTGGATTGGAAAGCAAGAAAA 1274

Qy        399 rIleThrPro------IleGlnTyrPheAsnPheAlaAsn---------LeuSerValPh 414
              ::::|||         ::::::     ||||||                ||||||:::
Db       1275 TCTAACAAGTTTACATCTCGAAACCAACAACTTCACTGGGACAATTCCACTCTCCCTGC 1334

Qy        414 eSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleTh 434
               ::::::         |||   |||:::||||||             :::        :::||
Db       1335 TAATCTTACAAAGCTGACAAGGCTCTATCTAGCGAATAATGCATTCGACGGTATTGTAAC 1394

Qy        434 rSerGlnTyrGlyValSerArgValIle------PheAsnThrSerAsnIleAsnAsnVa 452
              |         |||    :::   :::          |||   ||||||   ||||||::
Db       1395 ACCCAACTTGGGAAGGCTTCAACCACTGTTGGAGTTGAATCTTAGCAAC---AATAATCT 1451

Qy        452 lProGlySerLeuArgTyrGlu--------------------------------------- 459
              :    ||||||:::      |||
Db       1452 CCAGGGTAGCATACCTCCAGAGTTAGGTAACCTCAAACAATTGTACATACTAGATCTTTC 1511

Qy        460 ---------------------ValProAlaAsnLeu---------------------- 464
                                   :::|||     |||
Db       1512 ATCAAATAAATTCAGTGGTGCAATTCCTGAAACTTTAGGCCAATGTCAAAATTTAGTTAT 1571

Qy        464 ------------------------------------------------------------ 464

Db       1572 CCTTCAAATGGAACAAAATGTTCTTACAGGAAATATCCCAACCACTTTCACAAATCTAAA 1631

Qy        465 ------------------------------------------ProSerGlnThrIleLe 470
                                                        |||       :::||
Db       1632 TAGCTTGAGCCTTCTCAATCTTTCCCACAATGCTTTGTCGGGCCCCTTGCCAGCTGTTCT 1691

Qy        470 uSerGluLeuPro--------------------------------------------- 474
              |::::::|||
Db       1692 AAATGATCTAAAGCTTCTCAGACTATGTGGAGGAGCCATGACTTTGCGAATGCGTCCATG 1751

Qy        474 ------------------------------------------------------------ 474

Db       1752 TCCTGTTTCTCATAGAGCAATAAGCGTAAACAGATTAATCAAAATATTGATCCCAATATT 1811

Qy        475 ------------------------------------------GlyLysAspLysProAr 480
                                                        ||||||   :::   ||
Db       1812 TGGCTTCATGTCGCTCATATTATTGATCTACTTTTTACTCCTTGGGAAGAAGAGGTCAAG 1871

Qy        480 g-----------------ProAsnAlaGlyAspPheSerHisArgLeuSerTyr----- 492
              |                 |||:::   |||:::      ::::::||||||
Db       1872 AAGAACATCATTATCAGAGCCTTCTTTTGGTGAGCATTTTGAAAAAGTTTCTTATAATGA 1931

Qy        493 ---------------------IleSerAsnPheAsp---AlaArgArgSerSerSerGl 504
                                   |||   |||:::         |||   |||   ||
Db       1932 TCTTGCTCAAGCAACAAGGGACTTCTCGGAATTCAACCTCATTGGGAGGGGAAGCTACGG 1991

Qy        504 yGlyIle--------------------------------ValSerLeuLeuThrPheGl 513
              |   :::                                |||::::::     |||
Db       1992 TTCGGTATATCGAGGAAAATTAAAGGAAGCAAAATAGAAGTGGCTGTGAAGGTTTTCGA 2051

Qy        513 yTrpAlaHisThrSerMetAspArgAsn-----------------------AsnArgLe 525
                                  :::|||:::                          ::
Db       2052 CCTTGAGATGCATGGAGCAGAGAGAAGCTTCATGTCAGAATGTGAAGCACTAAGAAGCAT 2111

Qy        525 uGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGly-- 544
              ::::       :::         |||                     |||   |||   |||
Db       2112 TCAACATCGAAATCTTCTTTCTATCAGAACTGCTTGCTCAACAGGAGATAATACGGGCAG 2171

Qy        545 -------------PheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAs 560
                              |||                  |||||||||           ||
Db       2172 AGTTTTCAAAGCTTTAGTTTTTGATTACATGCCTAATGGGAACCTG-----------GA 2219

Qy        560 pSerTrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyr---ArgIleAr 579
              |   |||      |||      :::      ::::::      :::|||      :::
Db       2220 TATGTGG------CTACATCACAAAGAGGATGAGAAAACTCCGAATTATCTAAACTTAAC 2273

Qy        579 gLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIle---PheValGluHisSerGl 598
                 |||      :::      |||         ||||||:::   ::::::      |||
Db       2274 TCAGAGAATAAGCATAGCTGTTAACATTGCTGATGCTTTGGATTATTTGCACCACGACTG 2333

Qy        598 ySerSerHisIleValSerPheAspCysSerAsnSerSerGlyArgProSerAsnTh 618
                          ||||||                  :::|||      :::|||||||||
Db       2334 CGGAAGAACCATTGTG------------------CATAGTGATCTGAAGCCAAGCAACAT 2375

Qy        618 rLeuLeuGluSerAspPheArg-------------------------------TyrIl 627
                ||||||:::    |||                                    |||::
Db       2376 CCTTCTAGATGACGATATGAATGCTCTTCTGGGAGATTTTGGAATTGCACGTCTATATCT 2435

Qy        627 eAspValProGlyIlePheThrProSerIleAsn------------------------- 638
               :|||            :::      |||::::::
Db       2436 TGATTCTCAGTCAACGTACCAAGGCTCAGTTAGCACAATTGGTCTGAAGGGAACAATTGG 2495

Qy        639 -------ProLeuIleArgTyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPh 656
                        |||      :::      |||                |||      :::|||
Db       2496 TTATATTCCTCCAGGAAAAAGGCCAACAGATCCTATGTTCACGGACGGACTTGACATTGT 2555

Qy        656 eGluPheIleProLeuAsnThrPheProAsnGlnSerLeuGlu-----------LysAr 672
                   |||:::      |||   ||||||:::|||                      :::
Db       2556 CAGCTTTGTGGACAGAAAC---TTTCCGGATCAAATATTCCAAGTCATCGATGTTCGTCT 2612

Qy        672 gGluGlnGluValAsnAsp------------------LeuPheIleAsn 682
                ::::::|||   |||:::                   ||||||||||||
Db       2613 GAAGGAAGAATGCAACAACTTATCTCGAGATAACATGTTATTTATAAAT 2661
```

RESULT 7
CL853210

```
LOCUS       CL853210                 746 bp    DNA     linear   GSS 16-AUG-2004
DEFINITION  OR_CBa0082F07.r OR_CBa Oryza rufipogon genomic clone OR_CBa0082F07
            3', genomic survey sequence.
ACCESSION   CL853210
VERSION     CL853210.1  GI:51258277
KEYWORDS    GSS.
SOURCE      Oryza rufipogon
  ORGANISM  Oryza rufipogon
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 746)
  AUTHORS   Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M.,
            Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and
            Wing,R.
  TITLE     OMAP project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0082  row: F  column: 07
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..746
                     /organism="Oryza rufipogon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4529"
                     /clone="OR_CBa0082F07"
                     /tissue_type="young leaves"
                     /dev_stage="2 week old seedlings"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OR_CBa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;
                     drk treated 36 hrs before harvest"
ORIGIN

Alignment Scores:
```

```
Pred. No.:                 0.331          Length:          746
Score:                     108.00         Matches:         56
Percent Similarity:        36.2%          Conservative:    28
Best Local Similarity:     24.1%          Mismatches:      60
Query Match:               3.0%           Indels:          88
DB:                        10             Gaps:            11

US-10-782-096-2 (1-682) x CL853210 (1-746)

Qy        171 ArgAlaLeuValThrGln---------TyrAlaIleValAspAsnPhePheGlu------ 185
              :::|||:::::|||              |||      |||   ||||||::::::
Db         43 AAAGCCATTATCACCGATAAACTACTTTATCTGTACGTTAAAAATTTCTATGATTGAAAG 102

Qy        186 LysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyrAla 205
               |||            :::|||::::::    |||     :::    |||||||||||
Db        103 TGAAATTCCACCTGATATAAAAAAAAAACATTT---GTCTATCTTTTGCCAGTA------ 153

Qy        206 GlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrp 225
                           ||||||:::|||||||||           :::         |||       |||
Db        154 ---------AAATTGCATGTTATCTTGCTCTAGCGTCGTGAACAACATGGGATCCATTGG 204

Qy        226 GlnLeuGlyAspAspGluIleArgAspAsn---------TyrIleArgLeuGlnGlyLeu 242
              |||:::        |||                           :::|||   |||:::
Db        205 CAGATCTAGCGCGACAGTACCATGACTACCCCCGAGTAGCTGCCCAAACTTACCGGCATA 264

Qy        243 IleArgGluTyrLysAspHisCysIleThrPheTyr--AsnGlnGlyLeuAsnGlnPheA 262
                 :::|||   :::    ||||||         :::      |||    |||       |||    :
Db        265 TGGCAAGAGGTAGAAATACATTGTTATTGGTATCTGATAATGCCGGTGCATGTCAAGACT 324

Qy        262 snArgSer---------------------------------------------AsnAlaG 267
              ::|||                                                 :::|||
Db        325 CAAGAGGTAGAGAAATACCTCATTTGTTAGTACCAGGCAATGCAGACGACTTTAGCGCGG 384

Qy        267 lnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeuAsp---- 285
                 |||:::|||   |||    :::|||                        ||||||
Db        385 TGGATTTTGTTCTTTTCTAAAAATTC------------------------TTAGATCCAT 420

Qy        286 ---------------------------------------------------LeuAlaI 288
                                                                 |||||||
Db        421 CACTAGTGTGAAGGTGGTGGTAAGGGCTGAGATACATAAATTAGTAAGGGCTCCTAGCTA 480

Qy        288 leLeuPheProAsnTyrAspPro---ArgArgTyrProLeuAlaValLysThrGluLeuT 307
              |||||:::    ::::::    |||            ||||||   :::|||   ||||||
Db        481 TATTGTGGGCACACCATACACCTCTTCTGTCTTTGCCTCTGCTAATTAAAGATGAACTT- 539

Qy        307 hrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGlyGlyArgT 327

Db        539 ------------------------------------------------------------ 539

Qy        327 hrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgA 347
                                        :::|||         |||      ||||||    |||:
Db        540 ----------------------AGGAGTACTATTAAAGCTTACAATAATAATAAAAGAA 576

Qy        347 rgArgProSerTyrThrThrTrpLeuAsnArg 357
              ::    |||   :::         |||      |||
Db        577 AGAATCCACACTTT---GTATGGTCTACGAGG 605

RESULT 8
CL982610
LOCUS       CL982610              2439 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFSC048386 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL982610
VERSION     CL982610.1  GI:52419702
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 2439)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1..2439
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Alignment Scores:
Pred. No.:                 2.32           Length:          2439
Score:                     108.00         Matches:         145
Percent Similarity:        34.0%          Conservative:    108
Best Local Similarity:     19.5%          Mismatches:      255
Query Match:               3.0%           Indels:          240
DB:                        10             Gaps:            36

US-10-782-096-2 (1-682) x CL982610 (1-2439)

Qy          3 SerTyrLysAsnLysAsnGlu------------TyrGluMetLeuAspAlaLeuArgIle 18
              ||||||::::::         |||               |||:::::::::   :::
Db        436 AGTTACAGGCATGCCGGAGAAAACCTTGACATGTATCAATTGGTAGGTTCCGGAGGTTTC 495

Qy         19 AsnSerAsnMetSerAsnCysTyrProArg---------TyrProLeuAlaLysAspPro 35
              |||             :::   |||        |||         |||   :::
Db        496 AACCTTCCGTCAGGATATATTTCATCCGGGGAAATATCTATACCATGACGAATCAGAGC 555

Qy         36 GlnMetThrMetArgAsnThrAsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThr 55
                         :::   ||||||              :::|||                |||:::
Db        556 GATCCTATTGTCAACAATACC------------TATCTGACCTATCACAGGGAAAACAGT 603

Qy         56 GlnPhe---IleGlyAspIleSer----------------------------------- 62
              :::|||     ||||||:::|||:::
Db        604 GAATTCACCATTGGAAACATCAATAAATTACTGGAACTTTCTCTATTCGGAAGAGGGCTC 663

Qy         63 ------ThrTyrSerSerProGluAla--------------------------------- 69
                         |||:::||||||:::
Db        664 GAATATGCCTACACTTCTCCGGATAAAGACAAAAAAATTGAAGCAGGTTTTGTAGACCAG 723

Qy         70 ------AlaLeuSerValArgAspAlaValLeuThr---------GlyIleAsnSerVal 84
                    :::|||         |||::::::   |||                |||   :::
Db        724 ACCTACAGCCTGATCGAAAGAAATTCTTTTCTGAAATATGGTTATGGATTTTACACCAGA 783

Qy         85 GlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSer 104
              ||||||            ||||||         ||||||::::::
Db        784 GGAACC------------CTTGGGGCTCAAAACGCTTCCCGTAAT--------------- 816

Qy        105 ArgLeuIleGlyIleLeuTrpAlaGlyProAspProPheGluAlaLeuMet-ValLeuVa 124
                 :::   |||                     ||||||:::|||          ::::::
Db        817 ---ATCTCCGGAACGTATATTTTCAGGGATGATCCTTACGAAAAAGCCACACATCATGTA 873

Qy        124 lGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArg----- 142
                   |||   |||            :::      :::         |||        |||
Db        874 GCGGGAACAGATCTTCAGTATGCTTTCGGCAAAGACTGGAGAACCAACACCAAGGTGTAT 933

Qy        143 ------------------GluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnTh 156
                                ::::::      ||||||   :::        |||      :::::
Db        934 GGCGGACTAAGTTTCTACGAAAATAACCAGCCTGCAAAACCTTCCCTCGCTTTG-GAATC 992

Qy        156 rArgLeuGlnAlaTrpLeu------ValAsnLysAsnAspAsp---------------- 168
```

```
               ::::         :::      ||||||   |||   :::
Db        993 ACAATATTCCGGAATGATCAAAAAGGTCAATCTGAACGGAAATTATTTCTACAGCACAGA 1052

Qy        169 -------------AsnArgArgAlaLeuValThrGlnTyrAlaIleValAspAsnPhePh 184
                           |||||||||   ::::::            |||      |||||
Db       1053 TTATTATCCCGGAAACAGAAGAGGAATTTTG---------CAGATCCAGCAGAATTTC-- 1101

Qy        184 eGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTy 204
                  :::          ||||||:::  |||                        ||||
Db       1102 -------TCTACCATGATTTTCAAAGACCATAAT--------------------GTGTA 1133

Qy        204 rAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGl 224
              ||||                                      |||   :::|||      ::
Db       1134 TGCCAACATTACG--------------------------GCTTCCCATTTTTCACCCAA 1166

Qy        224 nTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeu-------------- 239
               ::::         |||:::   :::    :::   :::||||||
Db       1167 GTTCTACTTTTACGACAACACCCTGAATTCAAGCAATCTAAGACTGGATACGGGAATCAA 1226

Qy        240 ------------------------GlnGlyLeuIleArgGluTyrLysAspHisCysIl 251
                                        ||||||     :::|||::::::
Db       1227 CTTCCCTAAAAAAGGAAATTTCGGGATGGGACTTGGCTACCAGTATCAGGAAGAAAATTC 1286

Qy        251 eThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAsp-------- 268
                     ||||||        |||          |||   :::
Db       1287 CAATTCTTATAATAATTTTTTCAATGCCCAGCCTTATGAGAACACGAAACAGCTGAAAGC 1346

Qy        269 ------------------------TrpValSerPheAsnArgPheArgThrAspMetTh 280
                                      |||:::|||    ::::::      :::   :::
Db       1347 ACAGCGTTTTACGGAATACCTTACCTGGCTGAGCTCTGATAAGCAACACTCCTCTATC-- 1404

Qy        280 rLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLe 300
               |||::::::              :::   :::|||:::   |||   :::::::::
Db       1405 -TTAAGCATGGAGACAGGCTTGGTTCAATATCCGGACAACGATAAACAGCAGTTTCAGAC 1463

Qy        300 uAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyVa 320
                 |||
Db       1464 GAAAGTA----------------------------------------------------- 1470

Qy        320 lLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThr----------- 336
                      ||||||      :::|||   |||         |||
Db       1471 -------AGCGGAACCTACAGCTACAAATGGCTTACCGTGAACGGGATCTATCAGCATGG 1523

Qy        337 ---------------------ThrPheThrAlaMetGluAsnAsnAlaArg------Ar 347
                                   |||:::   |||   |||   :::
Db       1524 AAGTTATTTCCTTTCCGAATATGCTTTTCAAAAATGATGAACCAAAGCACTCCCTACGA 1583

Qy        347 gArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThrLeuGlyAs 367
                :::   |||   :::   ::::::|||:::   |||   :::   |||   |||
Db       1584 AAAGCTTTCCCTTTCGGCTTTTGTCAATAAAAATTTC---TTCAACCGTCAGCTG----- 1635

Qy        367 nMetSerAspValArgAsnIleTrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGl 387
                                                            |||       ||
Db       1636 ---------------------------------------------AATGTCACTTCAGG 1649

Qy        387 ySerGluIleThrHisAsn------PheGlyLysThrAspSer----------------- 399
              |         |||   :::       :::||||||:::   |||
Db       1650 GGTTTCTTATACGGATGATGTTCTGTACGGAAAATCGCCTTCCGGCTTTATGAACCTGAA 1709

Qy        400 -IleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLe 419
                 :::      :::|||     :::    |||    |||    ||||||
Db       1710 GTATTCAAGGGAAAAATATGCTCTGTACCTGAACTCATCCTACTTCAGT----------- 1758

Qy        419 uAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyVa 439
                         |||       |||:::        ||||||   :::|||   :::   |||||
Db       1759 ----------TATACTGCAGGAAGTTTCACCAATAATCTGTTGACCGTGGAAGCCGGTGT 1808

Qy        439 lSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGl 459
              |:::      :::   |||::::::                     |||   ::::::
Db       1809 AACCGTCAACCTTAGAAACAGCACCCTTGATCCGGGTAAAAAGGGTGATATCAAAGCATT 1868

Qy        459 uValProAlaAsnLeuProSerGlnThrIleLeuSerGluLeuProGlyLysAspLysPr 479
                |||       :::|||          |||       |||          ||||||

Db       1869 TGTCTATTATGACCTGAACGAAAATAATATCTATGATGAG---------GGCGACAAA-- 1917

Qy        479 oArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaAr 499
                                  :::   :::         |||      ::::::|||      ::
Db       1918 ----------------GAAGCGGCAGGCTATCTCATTATGCTCAACAATATTTCATTAA 1961

Qy        499 gArgSerSerSerGlyGlyIle-----------------------------------Va 507
                :         ||||||   |||
Db       1962 AACCGACCCATCAGGTTCTATCAGTTACCGTTCTATCCCTTATGGAAAGTATGCCCTGAA 2021

Qy        507 lSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluPr 527
                 ::::::         ||||||    :::              :::
Db       2022 ACAGGTCATCCAGCAAGGCTGGTATTACGATGAAACCGAA-----------TTTACCGT 2069

Qy        527 oAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPheValIl 547
                 ||||||                                   :::      |||    ||
Db       2070 AGACAAG------------------------------CACCACTACTCTTTTGAGAT 2096

Qy        547 ePro-----------GlyProThrGlyGlyAsnLeuValLysValSerAspSerTrpHi 563
                ||||           |||   |||   |||   :::                ||||||
Db       2097 CCCTCTTCATCAAAACGGAACGACCCAGGGAAAAATAACTTACGACTTTGATTCCAAAAC 2156

Qy        563 sSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAl 583
                ::::::   :::           |||:::   :::         |||   |||      |||
Db       2157 CGCAGTAG--ATTTCACGCCCAAAGTCGGAGGTGTTCTATTCAATATT-TACCGCAACGA 2213

Qy        583 aCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleVa 603
                  ||||||   |||           |||       :::     :::|||
Db       2214 ACAGTTAGTACATCAT---------ATCATCACCGATGATAACGGG------GAATTTGC 2258

Qy        603 lSerPhePheAspCysSer------------------AsnSerSerGlyArgProSerAs 617
                ||||||                                |||     :::      ||||||||
Db       2259 CTCCTTCCTGCCCTCCGGAAATTACAGGATTGAACTAAACAAAAACTCTCTGCCATCCAA 2318

Qy        617 nThrLeuLeuGlu 621
                ||||     |||
Db       2319 TACCTATTGCGAA 2331

RESULT 9
CL967463
LOCUS       CL967463                1437 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFCC039985 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL967463
VERSION     CL967463.1  GI:52389569
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1437)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1..1437
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
```

```
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Alignment Scores:
Pred. No.:              1.11          Length:         1437
Score:                  107.50        Matches:        78
Percent Similarity:     34.9%         Conservative:   55
Best Local Similarity:  20.5%         Mismatches:     89
Query Match:            3.0%          Indels:         159
DB:                     10            Gaps:           20

US-10-782-096-2 (1-682) x CL967463 (1-1437)

Qy        311 TyrThrAspProValGlyPhe------ThrGlyValLeuGluSerGlyGlyArgThrTyr 328
              :::   |||   :::|||              ||||||||  |||     :::|||
Db        112 TTCAAGGATGCCATCGGCGACGATCCGGCGGGCGTCCTGAGCTCCTGGAACAAAACCACC 171

Qy        329 ProTrpTyrAsnProAsnAsnThr--------------------ThrPheThrAlaMet 341
              |||:::         |||                           ||||||:::
Db        172 CCTTTCTGCCGGTGGAATGGCGTGAAATGCGGCCGCCGGGAGCATCGCGTCACGGCGCTC 231

Qy        342 Glu---------AsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
              |||         |||      |||
Db        232 GAGCTCGCCGGCCAGAACTTGACCGGTCGA------------------------------ 261

Qy        359 PheValTyrThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHis 378
                 :::         :::|||||||||:::|||   :::                  |||
Db        262 ---CTCGCCGCCGCTTCTCTCGGGAATCTGAGTTACCTT------------------CAC 300

Qy        379 ThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp 398
               |||      :::||||||               :::          :::    :::
Db        301 CTGCTTGACCTTTCTGGTAAT---------CGCTTCTCTGGGCAGATACCTCGCCTGAAC 351

Qy        399 SerIleThrProIleGlnTyrPheAsnPheAlaAsn------------------------ 410
              |||:::      :::|||      |||   :::|||
Db        352 AGTCTCCGAAAATTGCAAGTCCTCAACCTGAGTAACAATATTTGGACGGGATCATTCCG 411

Qy        410 ------------------------------------------------------------ 410

Db        412 GACACGCTTACAAACTGTTCCAGTTTGACGCAGCTAGATTTGAGCATAAACTTATTTCAA 471

Qy        411 --------------------------LeuSerValPheSerIleGluSerLeuAlaArg 421
                                        ||||||      |||      |||:::|||
Db        472 GGCCAGATTCCCCTCGGCATAGGACTCCTCTCTGAGCTGTCGGATCTCGTGCTTTCTCGT 531

Qy        422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGly--------- 438
                 ||||||   |||               :::|||   |||:::   |||
Db        532 AATTATCTCTCGGGG---------------CACATTCCTTCTGAGCTCGGGAAGTTGTCG 576

Qy        439 ------------------------------------------ValSerArgValIle 443
                                                        :::   |||   :::
Db        577 AAATTGTCGTCGTTGGACCTGAGTGTAAATATCATATCAGGTGAAATTCCACGCGCCCTT 636

Qy        444 PheAsnThrSerAsnIle---------------------------------AsnAsnVal 452
              :::|||   |||::::::                                  ::::::
Db        637 TATAATCTCTCTTCTCTTCGAATGCTATTCTTGGAGATGAATTCTCTGGGGAAATCGTTG 696

Qy        453 ProGlySerLeuArgTyrGluValPro--------------------------------- 461
              |||   ::::::   |||   :::|||
Db        697 CCGTCAAACATTGGCTATGCCCTTCCTAATCTCCAATGGCTCTTATTGGGTGATAACATG 756

Qy        462 ------AlaAsnLeuProSerGlnThrIleLeuSerGluLeuProGlyLysAspLysPro 479
                    |||:::|||:::
Db        757 TTTCAAGGTAACATACCAGCA--------------------------------------- 777

Qy        480 ArgProAsnAlaGlyAspPheSerHis---ArgLeuSerTyrIleSer------AsnPhe 496
                    :::   |||:::   |||          |||   |||:::|||      :::|||
Db        778 ------TCCCTTGGCAACATTTCACAGCTACATTTGATATATCTGTCAGAGAACGATTTC 831

Qy        497 AspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHis 516
                    |||      ||||||   |||   :::|||   |||:::            :::
Db        832 TCTGGGCGAATTCCTAGCTCTCTCGGTAAGCTTTCGAACCTGTCTGTCAACCTTCAGTAC 891

Qy        517 ThrSerMetAspArgAsnAsn---------------------------ArgLeuGlu 526
                    :::||||||||||||                              |||:::
Db        892 CTGCTTTTGGACAGAAACAATCTATCAGGACACATACCGTCCAACATGGAAATCTTCAA 951

Qy        527 ProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPheVal 546
                    ::::::||||||:::|||                         |||:::
Db        952 ------CAACTCACACAGCTGGACCTTAGT---------GACAATAATCTAAAAGGCAAG 996

Qy        547 IleProGlyProThrGlyGlyAsnLeuValLysValSerAsp---------SerTrpHis 563
              :::|||   |||:::   ||||||||   ::::::::          |||::::::
Db        997 ATGCCA---CCCAGCTTGGGAAATCTTCAACAACTCACACAGCTGGACCTTAGTTACAAT 1053

Qy        564 SerLeuLysValGlnAlaPro-----------------GlnArgGlnThrSerTyrArg 577
              :::||||||   :::   |||                 ||||||      |||:::
Db       1054 AATCTAAAAGGCAAGATGCCACCCAGCTTGGGAAATCTTCAGCGGCTAGTGAGCTTTAAT 1113

Qy        578 Ile 578
              :::
Db       1114 CTC 1116

RESULT 10
BH770523/c
LOCUS       BH770523                1926 bp    DNA     linear   GSS 01-MAY-2002
DEFINITION  LLMGtag287 MG1363 Random Sequence Tag Library Lactococcus lactis
            subsp. cremoris genomic, genomic survey sequence.
ACCESSION   BH770523
VERSION     BH770523.1  GI:20373480
KEYWORDS    GSS.
SOURCE      Lactococcus lactis subsp. cremoris
  ORGANISM  Lactococcus lactis subsp. cremoris
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Lactococcus.
REFERENCE   1  (bases 1 to 1926)
  AUTHORS   Bolotin,A., Ehrlich,S.D. and Sorokin,A.
  TITLE     Studies of genomes of dairy bacteria Lactococcus lactis
  JOURNAL   Sci. Aliments (2002) In press
COMMENT     Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            best homologue in strain IL1403 is arcD1 (94%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 1898.
FEATURES             Location/Qualifiers
     source          1..1926
                     /organism="Lactococcus lactis subsp. cremoris"
                     /mol_type="genomic DNA"
                     /strain="MG1363"
                     /sub_species="cremoris"
                     /db_xref="taxon:1359"
                     /clone_lib="MG1363 Random Sequence Tag Library"
                     /note="Vector: pSGMU2; Site_1: SmaI; Library of
                     chromosomal fragments of L.lactis strain MG1363 was
                     prepared by partial AluI digestion or by sonication."
ORIGIN

Alignment Scores:
Pred. No.:              1.8           Length:         1926
Score:                  107.50        Matches:        58
Percent Similarity:     36.4%         Conservative:   40
Best Local Similarity:  21.6%         Mismatches:     94
Query Match:            3.0%          Indels:         77
DB:                     9             Gaps:           10

US-10-782-096-2 (1-682) x BH770523 (1-1926)

Qy        375 TrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
               ||||||            ||||||      |||   :::|||:::          :::
Db        952 TGGGGATTGACTTTACTTGTTATGCGAGGTGTTGAAGGGGCTGCTGCCATTAATGCAATT 893
```

```
Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
                          :::  |||:::      |||      |||   :::      |||
Db        892 GTCCTTGTTGCAAAATTGATTCCCTTGTTTGTCTTTGTTATTGCTGGAATAGTTACTTTC 833

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
                                                                     :::
Db        832 AAAGCTGGA------------------------------------------GTCTTT 818

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
              |||   :::                 :::  |||||          |||      |||
Db        817 AGTGCTCACTTTTGGCAAAATTTCGTTGCGAATACA------------AATGCTGATGGC 770

Qy        455 SerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeuSerGluLeuPro 474
                 ::::::               :::|||:::         :::   |||::::::
Db        769 GTAATTAAGAGCTTAACTTGGTCTAATATGACTGGCGGAGATTTATTCAGTCAAGTCAAA 710

Qy        475 GlyLys-----------------------------AspLysProArgProAsnAlaGly 484
              |||                                   :::                |||
Db        709 GGTTCACTCATGGTTATGATTTGGGTCTTCGTCGGAATCGAAGGGGCTGCTATGATGGGT 650

Qy        485 AspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSerGly 504
              |||   :::   |||  |||                 ||||||   ::::::|||
Db        649 GACCGTGCCAAACGTAAATCG---------------GATGCTGGTAAAGCTTCAATTTTC 605

Qy        505 GlyIleVal--------------------------SerLeuLeuThrPheGlyTrpAla 515
              |||::::::                          |||||||||   ||||||:::
Db        604 GGTTTGATTGCCTTGTTAGTGATTTATATCTTGCTCTCATTGTTACCATTTGGTTTTATG 545

Qy        516 HisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAla 535
                 :::      :::         |||      :::|||      :::       ::::::|||
Db        544 ---AGCCAACAAGAACTAGCCAATACAGGTCAACCTGGTTTGGTTCATATTTGAACGCT 488

Qy        536 ---ValLysGlyTrpGlyGlyAsn---------IleGlyPheValIleProGlyProThr 551
                 |||   ||||||||||||:::            ||||||   ||||||
Db        487 ATGGTTGGCGGTTGGGGTGGTTCACTGATGGCCATTGGTCTTGTGATT----------- 440

Qy        552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
                    :::|||:::               ||||||      |||   |||:::|||   |||
Db        439 ------TCACTTCTTGGAGCTTGGTTGTCATGGACAATGCTTCCTGTTGAAGCAACACAA 386

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::   :::    ::::::
Db        385 CAATTATCAGAACAAAAATTA--------------------------------------- 365

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                                            :::   |||:::|||         |||
Db        364 ------------------------------CTTCCTAGTTGGTTTGGTAAACTTAATGAC 335

Qy        612 SerGlyArgProSerAsnThrLeuLeu 620
                 |||   |||   |||:::||||||
Db        334 AAAGGTGCCCCTAAAAATTCACTTTTG 308
```

RESULT 11
CO516871
LOCUS   CO516871   500 bp   mRNA   linear   EST 15-JUL-2004
DEFINITION s13dSG98H0600058_447070 Glandular trichomes Medicago sativa cDNA, mRNA sequence.
ACCESSION   CO516871
VERSION   CO516871.1 GI:50321745
KEYWORDS   EST.
SOURCE   Medicago sativa
  ORGANISM Medicago sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE   1 (bases 1 to 500)
  AUTHORS   Aziz,N., May,G.D., Paiva,N.L. and Dixon,R.A.
  TITLE   Alfalfa trichome Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research
  JOURNAL   Unpublished (2004)
COMMENT   Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 224 6650
            Fax: 580 224 6692
            Email: gdmay@noble.org.
FEATURES   Location/Qualifiers
     source   1..500
                     /organism="Medicago sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:3879"
                     /tissue_type="Glandular trichomes isolated from stem"
                     /dev_stage="Trichomes were removed from internodes of 8-12 inch tall stems"
                     /clone_lib="Glandular trichomes"
                     /note="Vector: pDNR-LIB; Glandular-haired alfalfa plants were established in a Conviron growth chamber (16-h days, full lights, 24oC set point) in 40 one gallon pots. Plants were grown in Metromix 350 and fertilized with MiracleGro as needed. They were cut back closely to encourage the emergence of vigorous shoots. Trichomes were isolated from stems, approximately 8-12 inches long, clipped from plants 2-3 inches above the crown. With minimal handling of the stem, the apical bud, leaves and nodes were discarded, and the trichomes isolated from the internode segments."
ORIGIN

Alignment Scores:

Pred. No.:	0.225	Length:	500
Score:	107.00	Matches:	44
Percent Similarity:	42.6%	Conservative:	25
Best Local Similarity:	27.2%	Mismatches:	54
Query Match:	3.0%	Indels:	39
DB:	7	Gaps:	8

US-10-782-096-2 (1-682) x CO516871 (1-500)

```
Qy        341 MetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheVal 360
              ||||||||||||         :::   |||:::      :::   ::::::::   |||
Db         20 ATGGAGAACAACCATAACAATCAACAATCTTTTGGCAATTCAGTGATCAACTTCGTGTC 79

Qy        361 TyrThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrp--------------- 375
              :::|||      |||   |||:::|||         :::||||||
Db         80 CACACATCAAATCTAGCAAACCTATCTCTTAATGATTCAATCTGGGGAAACAATTACTCA 139

Qy        376 --------------------------------------GlyGlyHisThrLeuValGlu 382
                                                    ||||||      ::::::
Db        140 TCCTCCAAAAATCAAAGGAGAAATTTTGATACAAAACTTGGTGGTGAAATCATCATCAAC 199

Qy        383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAspSerIleThrPro 402
              |||   |||:::         :::               ||||||         :::|||
Db        200 AACAACAACAACTCTGTTTCTCTT------------GGAAAA------------TCCCCT 235

Qy        403 Ile---GlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
              :::      :::||||||               |||   ||||||   :::
Db        236 GTTTCTGATCATTTCAATAATCAAGGATCATCTTTTTTCTCTTTTGATGTTGGTGTTTCT 295

Qy        422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
                       ||||||                  |||            ||||||   :::
Db        296 CCTATTAATGGAGGTTTTAATAAGGGAATTTATTCAAAGCCTTCTTATGGTGGGAAT--- 352

Qy        442 ValIlePheAsnThrSerAsnIleAsnAsnValProGly-SerLeuArgTyrGluValPr 461
                    ||||||   ||||||:::|||            |||      :::|||:::   :::
Db        353 ------TTTAAT---AGTAACATGAATGTTACTAAAGGTGAAGATGAGATTTTTCATCCA 403

Qy        461 oAlaAsnLeuProSer------GlnThrIleLeuSerGluLeuProGlyLysAspLysPr 479
               |||||||||            ||||||::::::                   |||:::|||
Db        404 TCTAAATCTTCCAAGAAGAATTCAAACCCTAATA---------------AAAAACAAGGG 448

Qy        479 oArg 480
               |||
Db        449 GAGA 452
```

```
RESULT 12
BC032293
LOCUS       BC032293                4979 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Mus musculus RIKEN cDNA D030051N19 gene, mRNA (cDNA clone
            IMAGE:5400613), with apparent retained intron.
ACCESSION   BC032293
VERSION     BC032293.1  GI:22749640
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4979)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 4979)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 65 Row: n Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 27369963
            This clone has the following problem: retained intron.
FEATURES             Location/Qualifiers
     source          1..4979
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5400613"
                     /tissue_type="Eye, retina, mouse strain C57Bl\6"
                     /clone_lib="NIH_MGC_94"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
Pred. No.:              9.82        Length:          4979
Score:                  107.00      Matches:         84
Percent Similarity:     37.0%       Conservative:    51
Best Local Similarity:  23.0%       Mismatches:      111
Query Match:            3.0%        Indels:          119
DB:                     4           Gaps:            20

US-10-782-096-2 (1-682) x BC032293 (1-4979)

Qy        304 ThrGluLeuThrArgGlu---------------ValTyrThrAspProValGlyPheThr 318
              |||||||||:::                    :::   :::  ||||||      |||
Db        936 ACAGAGTTATCTCATTACCGTCAACGTGCCCTCCTGCAATCACAGCCAGTTCGCCGGACA 995

Qy        319 GlyValLeu-----------------------------GluSerGlyGlyArgThrTyr 328
                 :::|||                              ||||||   :::|||
Db        996 CCTCTCCTCCACAATTTCCTGCACATGCTGTCCTCCCGTTCCTCTGGCATCCAGACCGAG 1055

Qy        329 ProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArg 348
              |||::::::   |||           :::   |||          :::           |||
Db       1056 CCCTTCCATCCCCCGGAGCAGGCTTCATCAACGCAGCAGGACCAGGGCCTCCTGAACCGG 1115

Qy        349 ProSer---TyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThrLeuGlyAsn 367
              ||||||   ::::::|||                  |||   :::   |||   ||||||
Db       1116 CCGTCTGCCTTCAGTACA------------------GTCCAGAGCAGCACTGCGGGCAAC 1157

Qy        368 -----------------MetSerAspValArgAsnIleTrpGlyGly------------ 377
                               :::         |||      ||||||
Db       1158 ACGCTCCGCAACCTCAGTCTGGGTCCTACCCGGCGCTCTTTGGGGGGCCCTTTGTCTAGC 1217

Qy        378 --------------HisThrLeuValGluAsnGlyAsnAspGlySerGluIleThr--- 391
                            |||      :::      |||      ||||||||   |||
Db       1218 CACCCTTCTAGGTATCACAGAGAACTAGCCCCTGGGCTGACAGGTTCTGAGTGGACCCGG 1277

Qy        392 --------------HisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPhe 406
                            ::::::   :::   :::|||:::   |||   :::
Db       1278 ACAGTACTCACTCTGAACTCTCGTTCTGAGGTAGAATCTATGCCCCCACCGAGG---ACC 1334

Qy        407 AsnPheAlaAsnLeuSerValPheSerIle------------------GluSerLeuAla 420
              :::   ::::::::|||:::   |||:::                   |||   |||
Db       1335 AGTGCCTCCTCAGTGAGTTTGCTTTCTGTACTGAGACAGCAGGAAGGTGGCTCTCAGGCA 1394

Qy        421 ArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSer 440
                 :::|||         ||||||         :::   :::|||
Db       1395 TCTGTGTATACTTCAGCCACAGAAGGGAGGGGTTTTCCATCATCAGGGTTGGCAACTGAG 1454

Qy        441 ArgValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluVal 460
                          |||   |||:::   ||||||      ||||||:::|||:::|||:::
Db       1455 TCAGATGGAGGGAATGGTTCCAGCCAAAACAAC---TCAGGCAGCATTCGCCATGAGCTT 1511

Qy        461 ProAlaAsnLeuPro------------------------------------------- 465
                 :::|||
Db       1512 CAATGTGACCTGAGACGCTTCTTTTTGGAGTATGACAGGCTGCAGGAGCTGGATCAGAGC 1571

Qy        466 --------------------------------------------SerGlnThrIleLeu 470
                                                          :::   |||
Db       1572 CTGAGTGGGGAAACTCCCCAGACCCAACAGGCCCAGGAAATGCTCAACAATAATATTGAA 1631

Qy        471 SerGluLeuProGlyLysAspLys---ProArgProAsnAlaGlyAspPheSerHisArg 489
              ||||||   ||||||               |||   |||::::::   :::   :::
Db       1632 TCTGAGAGACCAGGTCCTTCCCACCTGCCTACCCCACACAGCAGTGAGAACAACTCCAAC 1691

Qy        490 LeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeu 509
              ||||||         :::   :::      |||:::                  :::|||
Db       1692 CTGTCTCGG---GGCCACCTTAATCGATGCCGTGCTTGCCAT------------AATCTT 1736
```

```
Qy      510 LeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLys 529
            |||||||||                           |||:::  |||       ::::::
Db     1737 TTAACCTTC------------------------AACAATGATACCCTTCGTTGGGAAAGA 1772

Qy      530 IleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPheValIleProGly 549
               |||                                                |||
Db     1773 ACCACA------------------------------------------------CCTAAC 1784

Qy      550 ProThrGlyGlyAsnLeuValLysValSerAspSerTrpHis---Ser-LeuLysValGl 568
               :::    |||        :::   |||   ||||||||    ||| |||   :::::
Db     1785 TACTCATCTGGT---------GAGGCTAGCTCTTCCTGGCACGTCTCCACTACCTTTGAG 1835

Qy      568 nAla---ProGlnArgGlnThrSerTyrArgIle------ArgLeuArgTyrAlaCysLe 585
            :|||      |||      |||||||||   :::         ||||||      |||||
Db     1836 GGCATGCCACCAAGTGGCAACCAGCTACCACCTCTTGAGCGGACTGAGGGCCAAATGCCT 1895

Qy      585 uValThrHisGly 589
            |         |||
Db     1896 AGCTCCAGCAGGC 1908

RESULT 13
BI079460
LOCUS       BI079460               981 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  602873864F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5005694 5',
            mRNA sequence.
ACCESSION   BI079460
VERSION     BI079460.1  GI:14497790
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 981)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11045  row: p  column: 15
            High quality sequence stop: 764.
FEATURES             Location/Qualifiers
     source          1..981
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5005694"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam2"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

Alignment Scores:
Pred. No.:              1.17        Length:         981
Score:                  105.00      Matches:        93
Percent Similarity:     34.1%       Conservative:   46
Best Local Similarity:  22.8%       Mismatches:     125
Query Match:            2.9%        Indels:         144
DB:                     2           Gaps:           18

US-10-782-096-2 (1-682) x BI079460 (1-981)

Qy      283 ValLeuAspLeuAlaIleLeu---PheProAsnTyrAspProArgArgTyrProLeuAla 301
            |||||:::::::|||   |||   |||   :::   ||||||
Db        4 GTCCTGAACATGGCGGACCTCAGCTTTTCTGACGGCGATCCT---------------ACT 48

Qy      302 ValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeu 321
            |||:::|||   |||   |||                                 ||||||
Db       49 GTACGCACGCTTTTGCGGCGC---------------------------------GTGCTG 75

Qy      322 GluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMet 341
            |||:::                |||           :::
Db       76 GAGACGGCGGATTCGCGCACTCCGATGAGACGCCGGAGCACTCGGATC------------ 123

Qy      342 GluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyr 361
                  ||||||:::||||||   |||   |||   :::   ||||||
Db      124 ------AATGCACAAAGAAGACGGTCTCAAACGCCATACTCTAACAGGCAGGTAGCCAA 177

Qy      362 ThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHisThrLeuVal 381
            |||:::|||
Db      178 ACAAAGACAAGTGCTAGGAAGCAA------------------------------------ 201

Qy      382 GluAsnGlyAsnAspGlySerGluIleThr-HisAsnPheGlyLysThrAspSerIleTh 401
                     :::   |||:::   ::::::       :::      :::|||   |||
Db      202 ---------TCCCATGGAGCTAGGGTAAGTGTCCACCTACTCCATGAGTGACTTCATAGA 252

Qy      401 rProIleGlnTyrPheAsnPheAlaAsnLeuSer------------------------- 412
             |||         |||   |||   :::|||
Db      253 ACCCTAAAGTCTATTTTTGTTTTGAAGCCTACCTTGAGGGGCTCTCAAGCCAAAAGAGAG 312

Qy      413 -------------------------ValPheSerIleGluSerLeuAlaArgIleTyrLe 424
                                        :::   |||:::|||   :::|||   :::
Db      313 ACTAGTGGCGCTTGTACCCATAGTCAGTTGGCAGATCGACTCGTGTTCAAGGTCGTGGGC 372

Qy      424 u---GlyGlyThrGlu------------AlaAsnAsnTyrIleThrSerGlnTyrGlyVa 439
                |||||||||:::            |||      |||:::|||:::        ::
Db      373 GCCTGGAGGAACAGACACCCCGGACTCTGCTAAGGAATATCTTACTAACTGCCCCAGAAT 432

Qy      439 l-SerArgVal---------------------------------------IlePheAsnT 446
            : |||   |||                                       :::     :
Db      433 CGTCCACTGTGATGCCAGACCCAGTGGTGAAGCCAGCACAAGTACCAGAAGTGGCCCGGT 492

Qy      446 hrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProS 466
            ::|||         :::      |||||||||         :::|||        |||
Db      493 CCTCCAGACGGGAGAGCAGTCGGGGAAGCCTGGAGCTACACCTTCCTGAACTGGAGCCCC 552

Qy      466 erGlnThrIleLeuSerGluLeu------ProGlyLysAspLysProArgProAsnAlaG 484
                 |||:::         |||      |||||||||   |||   :::
Db      553 CCTCAACCCTGGCTCCGGGTCTAACAGCCCCTGGCAAAAGGAAGCAGAAACTGAGATTAT 612

Qy      484 lyAspPheSerHisArgLeuSerTyr---------IleSerAsnPheAspAlaArgArgS 501
                 |||         :::               :::|||      :::   ||||||
Db      613 CAGTATTTCAGCAGGAAGTGGACCAGGGGTTGCCTCTCTCCCAA---GAGCCACGTAGGA 669

Qy      501 erSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspA 521
            ||   |||      :::   ||||||   :::
Db      670 GTAGGAGTGCTGATGTCTCATCCTTGGCTAGC---------------------------- 701

Qy      521 rgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyG 541

Db      702 --------------------------------------------------------TCCT 705

Qy      541 lyAsnIleGlyPheValIleProGly---------ProThrGlyGlyAsnLeuValLysV 558
              |||:::   ||||||:::||||||            :::|||   |||||||||   :
Db      706 TCAATCTTACCTTTGTTCTACCTGGTCAGGCCAGAGACAGTGGAGAGAACCTGGTCTGGC 765

Qy      558 alSerAspSerTrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgI 578
            ::      :::   ||||||
Db      766 TCCGAAAGACGTCCCATTCC---------------------------------------- 785

Qy      578 leArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerG 598
                                   |||            |||   |||            |
```

```
Db        786 --------------GGCAGCTTTGTGAAATGGCTGGTGGCTCCTTTTGGCAGGGATCCTG 831

Qy        598 lySerSerHisIleValSerPhePheAspCysSerAsnSerSerGly-----------A 614
              ||   :::   |||:::                     :::||||||         |
Db        832 GAAAGAACAAGATCCTT-----------------AAGGCTTCTGGTCTTCCAGGTGACA 873

Qy        614 rgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspVal-----------P 630
              |||||   ||||||||:::   :::   :::|||  |||  |||
Db        874 GGCCACGAAACACCTTGGTGGCTGCCTCCTGGCGATGAATTGTGGTTTGGGAACACACAC 933

Qy        630 roGlyIlePheThrProSer 636
                 |||     |||||||||
Db        934 GAGGCTTTCTCACACCATCC 953

RESULT 14
BM904226
LOCUS       BM904226                1016 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6693004 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5499692
            5', mRNA sequence.
ACCESSION   BM904226
VERSION     BM904226.1  GI:19354603
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1016)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12133  row: g  column: 21
            High quality sequence stop: 659.
FEATURES             Location/Qualifiers
     source          1..1016
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5499692"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_85"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.23          Length:        1016
Score:                  105.00        Matches:       61
Percent Similarity:     34.5%         Conservative:  38
Best Local Similarity:  21.3%         Mismatches:    86
Query Match:            2.9%          Indels:        102
DB:                     3             Gaps:          13

US-10-782-096-2 (1-682) x BM904226 (1-1016)

Qy        304 ThrGluLeuThrArgGlu---------------ValTyrThrAspProValGlyPheThr 318
              |||||||||:::                     :::   :::   ||||||     |||
Db         70 ACAGAATTATCCCACTACCGTCAGCGTGCCCTCCTGCAATCACAGCCAGTTCGCCGGACG 129

Qy        319 GlyValLeu-----------------------------GluSerGlyGlyArgThrTyr 328
                 :::|||                                |||||   :::|||
Db        130 CCTCTCCTCCACAATTTCCTGCACATGCTGTCCTCCCGCTCTTCTGGCATCCAGACCGAG 189

Qy        329 ProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArg 348
              |||::::::   |||        :::   |||         :::              |||
Db        190 CCCTTCCATCCCCCGGAGCAGGCCTCGTCAACGCAGCAGGACCAGGGCCTCCTGAACCGG 249

Qy        349 ProSer---TyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThrLeuGlyAsn 367
              ||||||   ::::::|||                  |||   :::   |||   ||||||
Db        250 CCGTCTGCCTTCAGTACA------------------GTCCAGAGCAGCACTGCCGGCAAC 291

Qy        368 ------------------MetSerAspValArgAsnIleTrpGlyGly------------ 377
                                :::      |||      ||||||
Db        292 ACGCTCCGCAACCTCAGTCTGGGTCCTACCCGCCGCTCTTTGGGAGGGCCTCTGTCTAGC 351

Qy        378 ---------------HisThrLeuValGluAsnGlyAsnAspGlySerGluIleThr--- 391
                             |||      :::      |||      ||||||||   |||
Db        352 CACCCTTCTAGGTATCACCGAGAAATAGCTCCTGGGTTGACAGGATCTGAGTGGACCCGG 411

Qy        392 ---------------HisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPhe 406
                             ::::::      :::   :::|||:::   |||   :::
Db        412 ACAGTACTCAGTCTGAACTCCCGCTCTGAGGCGGAATCCATGCCCCCGCCCAGA------ 465

Qy        407 AsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGly 426
                    :::   |||      |||:::   |||:::   |||            ||||||
Db        466 ------ACCAGTGCCTCTTCGGTGAGTTTGCTGTCTGTGCTGAGACAGCAGGAAGGTGGC 519

Qy        427 ThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgValIlePheAsnThr 446
              ::::::|||:::      |||   ::::::      |||
Db        520 TCTCAGGCATCTGTGTACACTTCAGCCACAGAAGGG------------------------ 555

Qy        447 SerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSer 466
                                                                   |||:::
Db        556 ---------------------------------------------AGGGGTTTTCCGGCA 570

Qy        467 GlnThrIleLeuSerGluLeuPro---GlyLysAspLysProArgProAsnAlaGlyAsp 485
                    :::   :::|||         |||               :::   |||:::|||
Db        571 TCAGGGTTGGCAACTGAGTCAGATGGAGGGAATGGCTCCAGCCAAAACAACTCGGGCAGC 630

Qy        486 PheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGly 505
                    |||   |||               |||   ||||||
Db        631 ATTCGCCATGAGCTT------------CAGTGTGACCTGAGACGCTTCTTT--------- 669

Qy        506 IleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeu 525
                                                                      |||
Db        670 ---------------------------------------------------------CTG 672

Qy        526 GluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPhe 545
              |||   |||::::::   ::::::|||            ||||||   :::
Db        673 GAGTATGACCGGCTTCAAGAGCTGGATCAGAGCCTTGAGTGGGGAAGCTCC--------- 723

Qy        546 ValIleProGlyProThrGly 552
                    |||   |||||||||
Db        724 ------CCAGACCCAACAGGG 738

RESULT 15
CL961241
LOCUS       CL961241                3180 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFCC005799 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL961241
VERSION     CL961241.1  GI:52377209
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 3180)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
```

```
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1. .3180
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Alignment Scores:
Pred. No.:              8.06        Length:         3180
Score:                  105.00      Matches:        96
Percent Similarity:     35.8%       Conservative:   85
Best Local Similarity:  19.0%       Mismatches:     185
Query Match:            2.9%        Indels:         140
DB:                     10          Gaps:           23

US-10-782-096-2 (1-682) x CL961241 (1-3180)

Qy        117 PheGluAla---LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArg 135
              |||||||||   |||:::          :::|||||||||:::          :::
Db        610 TTTGAAGCAAAGTTACTTGAGTTTGAACAGGAGCTCATTAGGGCTGGTGCTGAGAATGAT 669

Qy        136 ValArgGluAsnAlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGln 155
                             :::|||   |||
Db        670 GCACTCTCAAGGTCACTCGAGGAA------------------------------------ 693

Qy        156 ThrArgLeuGlnAlaTrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThr 175
                       :::         |||:::   |||   |||:::   :::
Db        694 ---------CGGGGAGACCTTTTGATGAAGATTGATGAGGAAAAGGCTCGAGCAGAGGCT 744

Qy        176 GlnTyrAlaIleValAspAsnPhePheGluLysAsnMetPro---LysPheLysGluArg 194
              :::   ::::::   |||      :::      :::      :::   :::|||:::
Db        745 GAAATCGAAGTCTTGAAAAACACAATTCAGTCAGAGCTTGAAATCCGCAATGAGGAAAAG 804

Qy        195 AsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeu 214
              |||      :::            |||         |||         |||:::
Db        805 AACATGAGTGTGCGCTCAGCTGATGTT---------GCAACTAAACAGCACATGGAGGAT 855

Qy        215 LeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAsp 234
              ::::::                  |||:::   |||
Db        856 GTAAAGAAAATATCAAAACTGGAAGCTGAATGCCAA------------------------ 891

Qy        235 AsnTyrIleArgLeuGlnGlyLeuIle--------------------------------- 243
                       ||||||:::||||||:::
Db        892 ---------AGATTACGGGGTCTTGTTCGAAAGAAGTTACCAGGGCCTGCTGCATTAGCT 942

Qy        244 ------------------------------ArgGluTyrLysAspHisCysIleThrPhe 253
                                            |||:::|||   :::      :::
Db        943 CAAATGAAAATGGAAGTGGAGAGCTTGGGCAGAGATTATGGAGAAAGCAGGTTGCGACGA 1002

Qy        254 TyrAsnGlnGlyLeuAsnGlnPheAsnArg------------------------------ 263
                       :::   |||:::|||
Db       1003 TCGCCTGCAAAGAATTCCAGCTTCCATCGTCCTATGTCCCCTATGTCCCCTGTTCCTGAT 1062

Qy        264 ---------SerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThr 280
                       |||   |||            |||      |||   |||
Db       1063 TATGCATTTGAAAACTTACAGCACATGCAGAAAGAGAATGAGTTTCTGACTGCACGTTTG 1122

Qy        281 LeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeu 300
              |||::::::   |||         :::                  :::|||
Db       1123 TTATCCATGGAAGATGAAACCAAAATGCTAAAAGAGGCCTTGGCCAAAAGGAATAGT--- 1179

Qy        301 AlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyVal 320
                    :::   :::   :::|||   :::|||               |||
Db       1180 ------GAGCTGCAAACATCAAGAAACATGTATGCTAAGACAGCAGGCAAACTCCGTGGC 1233

Qy        321 LeuGlu-----------SerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThr 336
              ||||||           :::|||   :::            ||||||||
Db       1234 TTGGAAGTTCAAATGTTGACTGGTAGCCAACGTAAGAGTACTTCGAATCCCAACATGGAT 1293

Qy        337 ThrPhe------ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThr------ 352
                          |||:::      |||            ||||||   |||
Db       1294 ATCCACTTTGATGGTGCACTGAGTCAAAATGGAAGCAACCCACCTAGTATGACCTCCATG 1353

Qy        353 ------------------------------------ThrTrpLeuAsnArgIlePhe 359
                                                  :::|||   |||   :::
Db       1354 TCTGAAGATGGTGTTGATGATGAAGGAAGTTGCACTGAATCTTGGGCCAATGCTCTG--- 1410

Qy        360 ValTyrThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHisThr 379
                             :::      :::|||   ::::::         |||   :::
Db       1411 ---------------GTCTCTGAGCTCTCACACATCAAGAAAGAGAAAGGAGCTAAGAGC 1455

Qy        380 LeuValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAspSer 399
                 |||      |||::::::      |||:::      :::|||   :::   :::
Db       1456 AGTGTGACAGAAGGCTCCAATCGGTTGGAACTCATGGATGACTTCCTAGAGATGGAGAAA 1515

Qy        400 IleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeu 419
              :::   :::         :::   ||||||      |||   ::::::|||
Db       1516 CTAGCCTGTTTGTCA------TCTGAAGCAAATGGACATGTCAGTACTGTTGAA------ 1563

Qy        420 AlaArgIleTyrLeuGlyGlyThrGluAlaAsn---AsnTyrIleThrSerGlnTyrGly 438
                 ::::::   :::      |||||||||:::   :::   ||||||   :::   |||
Db       1564 ---AAAATGAAGATTGATGATACGGAAGCTTCTTTGTCTGGTATTACTGAAAGAGATGGT 1620

Qy        439 ValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyr 458
              |||                  ::::::   ::::::      :::||||||:::
Db       1621 GTTAAA------------GATTCGCAGTCAGTTTTGGCATTGCCAGGGACT--------- 1659

Qy        459 GluValProAlaAsn-----------------------------LeuProSerGlnThr 468
                    |||:::|||                           |||      |||:::
Db       1660 ------CCATCTAATAAACTGCAGCTGTCTGACAGCTCTCCACTCTTGAAACTACAGTCT 1713

Qy        469 IleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHis 488
                 :::|||   |||         :::   |||:::   |||||||||:::
Db       1714 AGAATATCTTCCTTGCTTGATTCTGAATCACCACAGAACAATGCTGGAAATATA------ 1767

Qy        489 ArgLeuSerTyrIleSerAsnPhe---------------AspAlaArgArgSerSerSer 503
                 |||   |||   |||                     |||         ||||||:::
Db       1768 ---CTAGATAGTATCAGAAATATTCTAAAGGATATTGAAGATGAGGCAGATTCATCAAAT 1824

Qy        504 -----------------GlyGlyIleValSerLeuLeuThrPheGly------TrpAla 515
                               |||   :::|||   :::            |||
Db       1825 GATAGCAAGACTCATCATGGTGACATGGTTGAAGTAGCTGACAATGGATCACTAATGAAG 1884

Qy        516 HisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIle------ 533
              |||:::|||         :::   ::::::   :::   :::         |||
Db       1885 CATTCAAGCAGCGGGAGCAAGCATGCTATGGATCAAGAACTAGTCAATGCCATTTTAAAG 1944

Qy        534 ------AspAlaValLys 537
                    |||   ||||||
Db       1945 ATTCAAGACTTCGTTAAG 1962

Search completed: January 21, 2006, 03:54:26
Job time : 7102 secs
```

```
                    GenCore version 5.1.7
         Copyright (c) 1993 - 2006  Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 20, 2006, 22:58:38 ; Search time 326 Seconds
                                            (without alignments)
                                            3718.703 Million cell updates/sec

Title:          US-10-782-096-2
Perfect score:  3597
Sequence:       1 MNSYKNKNEYEMLDALRINS..........TFPNQSLEKREQEVNDLFIN 682

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters:      2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10782096/runat_20012006_095215_24416/app_query.fasta_1.839
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10782096_@CGN_1_1_193_@runat_20012006_095215_24416 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/1/ina/1_COMB.seq:*
                2:  /cgn2_6/ptodata/1/ina/5_COMB.seq:*
                3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/1/ina/H_COMB.seq:*
                6:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                7:  /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
                8:  /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
                9:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                      Description
----------------------------------------------------------------------------
    1   942.5    26.2   3507  2  US-08-315-468-3          Sequence 3, Appli
    2   940.5    26.1   3504  3  US-10-089-678-2          Sequence 2, Appli
    3   940.5    26.1   3690  3  US-10-089-678-3          Sequence 3, Appli
    4   910.5    25.3   4173  3  US-09-661-322A-37        Sequence 37, Appl
    5   901.5    25.1   3471  2  US-07-876-280-29         Sequence 29, Appl
    6   901.5    25.1   3471  2  US-07-812-180A-1         Sequence 1, Appli
    7   901.5    25.1   3471  2  US-08-315-468-1          Sequence 1, Appli
    8   901.5    25.1   3471  3  US-07-941-650A-1         Sequence 1, Appli
    9     897    24.9   3684  2  US-08-448-170-7          Sequence 7, Appli
   10     897    24.9   3684  3  US-08-961-803-5          Sequence 5, Appli
   11     894    24.9   3684  3  US-09-661-322A-62        Sequence 62, Appl
   12     892    24.8   8854  3  US-09-053-549-1          Sequence 1, Appli
   13     889    24.7   3558  3  US-09-178-252-22         Sequence 22, Appl
   14     889    24.7   3558  3  US-09-826-660-22         Sequence 22, Appl
   15   886.5    24.6   4074  2  US-08-377-690-1          Sequence 1, Appli
   16     880    24.5   3624  2  US-07-951-715A-6         Sequence 6, Appli
   17     880    24.5   3624  2  US-08-459-448A-6         Sequence 6, Appli
   18     880    24.5   3624  3  US-08-459-595A-6         Sequence 6, Appli
   19     880    24.5   3624  3  US-08-459-504B-6         Sequence 6, Appli
   20     880    24.5   3624  3  US-08-459-444-6          Sequence 6, Appli
   21     880    24.5   3624  3  US-09-053-549-7          Sequence 7, Appli
   22     880    24.5   3624  3  US-09-547-422-6          Sequence 6, Appli
   23     880    24.5   3624  3  US-09-988-462-6          Sequence 6, Appli
   24     872    24.2   1929  3  US-09-178-252-24         Sequence 24, Appl
   25     872    24.2   1929  3  US-09-826-660-24         Sequence 24, Appl
   26   868.5    24.1   2133  3  US-09-661-322A-41        Sequence 41, Appl
   27     855    23.8   1953  2  US-08-315-468-5          Sequence 5, Appli
   28     855    23.8   1953  3  US-07-941-650A-3         Sequence 3, Appli
   29   849.5    23.6   2159  3  US-08-286-870A-7         Sequence 7, Appli
   30     843    23.4   3934  2  US-08-100-709-3          Sequence 3, Appli
   31     843    23.4   3934  2  US-08-176-865-3          Sequence 3, Appli
   32     843    23.4   3934  2  US-08-474-038-3          Sequence 3, Appli
   33     843    23.4   3934  2  US-08-779-046-3          Sequence 3, Appli
   34     843    23.4   3934  2  US-08-881-340-3          Sequence 3, Appli
   35   842.5    23.4   1959  3  US-09-661-322A-5         Sequence 5, Appli
   36   841.5    23.4   2430  6  PCT-US92-00040-1         Sequence 1, Appli
   37   840.5    23.4   1956  3  US-08-996-441B-51        Sequence 51, Appl
   38   840.5    23.4   1956  3  US-08-993-722A-51        Sequence 51, Appl
   39   840.5    23.4   1956  3  US-08-993-170A-51        Sequence 51, Appl
   40   840.5    23.4   1956  3  US-08-993-775B-51        Sequence 51, Appl
   41   840.5    23.4   1956  3  US-09-427-770-51         Sequence 51, Appl
   42   840.5    23.4   1956  3  US-09-427-769-51         Sequence 51, Appl
   43   840.5    23.4   2965  2  US-08-460-570-1          Sequence 1, Appli
   44   840.5    23.4   2965  2  US-08-460-570-2          Sequence 2, Appli
   45   840.5    23.4   2965  3  US-08-286-870A-1         Sequence 1, Appli

                              ALIGNMENTS

RESULT 1
US-08-315-468-3
; Sequence 3, Application US/08315468
; Patent No. 5554534
;  GENERAL INFORMATION:
;    APPLICANT:  Michaels, Tracy Ellis
;    APPLICANT:  Foncerrada, Luis
;    APPLICANT:  Narva, Kenneth E.
;    TITLE OF INVENTION:  Process for Controlling Scarab Pests
;    TITLE OF INVENTION:  with Bacillus thuringiensis Isolates
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David R. Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/315,468
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/014,941
;      FILING DATE:  01 FEB 1993
;      APPLICATION NUMBER:  07/828,430
;      FILING DATE:  30-JAN-1992
;    PRIOR APPLICATION DATA:
```

```
;     APPLICATION NUMBER:  07/808,316
;     FILING DATE:  16-DEC-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Saliwanchik, David R.
;     REGISTRATION NUMBER:  31,794
;     REFERENCE/DOCKET NUMBER:  MA73.C2
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  904-375-8100
;     TELEFAX:  904-372-5800
; INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  3507 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   ORIGINAL SOURCE:
;     ORGANISM:  Bacillus thuringiensis
;     STRAIN:  kumamotoensis
;     INDIVIDUAL ISOLATE:  50C
;   IMMEDIATE SOURCE:
;     LIBRARY:  LambdaGEM-11(tm) library of L. Foncerrada
;     CLONE:  50C(b)
US-08-315-468-3

Alignment Scores:
Pred. No.:              3.52e-90      Length:         3507
Score:                  942.50        Matches:        249
Percent Similarity:     50.3%         Conservative:   115
Best Local Similarity:  34.4%         Mismatches:     272
Query Match:            26.2%         Indels:         87
DB:                     2             Gaps:           27

US-10-782-096-2 (1-682) x US-08-315-468-3 (1-3507)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGTTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAGGTATTTCTAAGCGAGCAAGATGCAGTTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTAACAGGTTTAGGGGTTCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGly------ProAspProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAAAACAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTGGAAGGGCTAGGGAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAGTGGAAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTTCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTCGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACAGTATATACAATGGCAGCAAACCTACATTTACTTTTATTAAGGGACGCA 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATTGTCTACAAGCACTATTAATAACTACTATAATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGACCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCTCGAGCGCTAAACAATGGATTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACATTGACGGTGTTAGACGTTGTTGCATTATTTTCAAACTATGATACGCGTACGTAT 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCACTGGCAACAACAGCTCAGCTTACAAGGGAAGTATATACAGATCCACTTGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 GATGTGCCTAATATTGGCTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 TCAGAAATAGAAAAAGCGGCTATTCGTCCACCTCATGTGTTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThrArgThrLeuGlyAsnMetSerAsp---ValArgAsnIleTrpGlyGly 377
Db       1048 ACAGTTTATACAAAAAAACGTAGCTTCACTTCTGATCGTTATATGAGATATTGGGCTGGT 1107

Qy        378 HisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLys 396
Db       1108 CATCAAATAAGCTATAAGCATATCGGTACGAGTAGTACCTTTACACAGATGTATGGAACC 1167

Qy        397 ThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIle 416
Db       1168 AATCAAAATTTACAAAGTACTAGCAATTTTGATTTTACGAATTACGATATTTACAAGACT 1227

Qy        417 GluSer--------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsn 431
Db       1228 TTATCAAATGGTGCAGTACTCCTTGATATAGTTTACCCTGGT----------------- 1269

Qy        432 TyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIleAsn 450
Db       1270 TATACGTATACATTTTTTGGAATGCCAGAAACCGAGTTTTTTATGGTAAATCAATTGAAT 1329

Qy        451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
Db       1330 AATACCAGAAAGACGTTAACGTATAAA---CCAGCT---------TCCAAAGATATTATA 1377

Qy        471 Ser-------------------GluLeuProGlyLysAspLysProArgProAsnAla 483
Db       1378 GATCGGACAAGAGATTCGGAATTAGAATTGCCTCCAGAAACTTCAGGTCAACCAAATTAC 1437

Qy        484 GlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSer 503
Db       1438 GAGTCATATAGCCATAGATTAGGTCATATTACATTTATT------------TACTCCAGT 1485

Qy        504 GlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsn 523
Db       1486 TCAACTAGCACGTATGTACCTGTATTTTCTTGGACACATCGGAGTGCAGATCTAACAAAT 1545

Qy        524 ArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIle 543
```

```
Db       1546 ACAGTTAAAAGTGGCGAAATCACCCAAATACCAGGG------GGCAAGTCTAGCACCATA 1599

Qy        544 Gly---------PheValIleProGlyPro------ThrGlyGlyAsnLeuValLysVal 558

Db       1600 GGCAGAAATACTTATATAATAAAAGGGCGTGGTTATACAGGGGGAGACTTAGTGGCTTTA 1659

Qy        559 SerAspSerTrpHisSerLeuLysValGlnAla------ProGlnArgGlnThrSerTyr 576

Db       1660 ACGGACCGCATCGGAAGTTGTGAGTTTCAGATGATCTTTCCAGAGTCTCAA---CGATTC 1716

Qy        577 ArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHis 596

Db       1717 CGTATTCGGATTCGTTACGCT--------------------------------------- 1737

Qy        597 SerGlySerSerHisIleValSerPhePheAspCysSerAsnSer--------------- 611

Db       1738 TCTAATGAAACTAGTTATATTAGTTTATACGGACTAAACCAAAGCGGAACTTTAAAATTC 1797

Qy        612 ------------SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIle 627

Db       1798 AACCAGACATATTCTAATAAAAATGAAAATGATTTAACATATAATGATTTCAAATATATA 1857

Qy        628 AspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArgThrGlnSer 647

Db       1858 GAATATCCAAGAGTCATT------TCAGTAAATGCTTCTTCAAACATACAGAGGTTATCT 1911

Qy        648 PheGlyThrHisAla------------------IleAspLysPheGluPheIleProLeu 661

Db       1912 ATAGGTATACAAACGAATACAAATTTATTTATTTTAGACCGAATCGAATTCATCCCAGTA 1971

Qy        662 Asn---ThrPheProAsnGlnSer---LeuGluLysArgGluGlnGluValAsnAspLeu 679

Db       1972 GATGAGACATATGAAGCGGAAACGGATTTAGAAGCGGCAAAGAAAGCAGTGAATGCCTTG 2031

Qy        680 PheIleAsn 682

Db       2032 TTTACGAAT 2040
```

RESULT 2
US-10-089-678-2
; Sequence 2, Application US/10089678
; Patent No. 6962977
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A
; TITLE OF INVENTION: NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD
; FILE REFERENCE: Q68821
; CURRENT APPLICATION NUMBER: US/10/089,678
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3504
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(3501)
; OTHER INFORMATION:
US-10-089-678-2

Alignment Scores:

Pred. No.:	5.76e-90	Length:	3504
Score:	940.50	Matches:	256
Percent Similarity:	51.8%	Conservative:	111
Best Local Similarity:	36.1%	Mismatches:	288
Query Match:	26.1%	Indels:	56
DB:	3	Gaps:	23

US-10-782-096-2 (1-682) x US-10-089-678-2 (1-3504)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22

Db          4 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42

Db         61 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61

Db        121 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81

Db        181 GAGACGTTTATTAGT--------------------TCATCTACGGTTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101

Db        220 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119

Db        280 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139

Db        340 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159

Db        400 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 459

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178

Db        460 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 519

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198

Db        520 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 579

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218

Db        580 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 639

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238

Db        640 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 699

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258

Db        700 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 759

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278

Db        760 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 819

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298

Db        820 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 879

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318

Db        880 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 939

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338

Db        940 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTCTTTT 996

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
```

```
Db        997 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTGATATACTCAGTGAAATA 1056

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375
                   ||||||          |||          |||   ::::::            |||
Db       1057 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1116

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
                 ||||||::::::      :::|||   |||     |||    :::     |||:::
Db       1117 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1176

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
              |||   |||   :::          |||:::|||::::::   ||||||     :::|||
Db       1177 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1233

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
                 :::   ||||||                   |||    :::     ||||||
Db       1234 CAGGTTCGATCATTA------------------GGGGCGGATTTAGCTAATTACTACGCA 1275

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
                    |||||||||          |||          :::   |||            |||
Db       1276 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1335

Qy        455 SerLeu---------ArgTyrGluValProAlaAsn---------------LeuProSer 466
              |||:::          |||   |||                                :::
Db       1336 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1395

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
                    :::   |||:::|||   :::::::::|||                      :::
Db       1396 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1449

Qy        487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504
              ||||||||||||||||||:::|||:::           :::|||      ||||||
Db       1450 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTTCTAAGAATGCTAGTAGTCCTGCT 1509

Qy        505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524
                    :::|||       |||   |||   |||   |||   |||         |||
Db       1510 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1569

Qy        525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541
              :::      ||||||||||||||||||         ||||||                |||
Db       1570 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1629

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
                                ||||||   ||||||||||||::::::|||   :::    |||
Db       1630 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1689

Qy        562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577
                          :::   |||:::    |||:::::|||   |||         ||||||
Db       1690 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1746

Qy        578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597
                 |||:::||||||||         |||          :::||||||   |||
Db       1747 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1800

Qy        598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617
              |||   ||||||::::::       :::   |||   |||   :::            :::
Db       1801 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGGAT 1851

Qy        618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636
                 |||   :::   |||               :::         ||||||    :::
Db       1852 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 1911

Qy        637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653
                 :::      :::               ||||||          |||         :::
Db       1912 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 1971

Qy        654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671
              ||||||||||||   ||||||:::|||          ||||||    :::   |||:::
Db       1972 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2031

Qy        672 ArgGluGlnGluValAsnAspLeuPhe 680
                 ::::::   ||||||   ||||||
Db       2032 GCAAAGAAAGCAGTAAATGGCTTGTTT 2058
```

```
RESULT 3
US-10-089-678-3
; Sequence 3, Application US/10089678
; Patent No. 6962977
; GENERAL INFORMATION:
;  APPLICANT: ASANO, Shinichiro
;  TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A
;  TITLE OF INVENTION:  NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD
;  FILE REFERENCE: Q68821
;  CURRENT APPLICATION NUMBER: US/10/089,678
;  CURRENT FILING DATE:  2002-04-03
;  PRIOR APPLICATION NUMBER: JP 2000-236140
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: PCT/JP01/06660
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 3690
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-10-089-678-3

Alignment Scores:
Pred. No.:              6.26e-90       Length:        3690
Score:                  940.50         Matches:       256
Percent Similarity:     51.8%          Conservative:  111
Best Local Similarity:  36.1%          Mismatches:    288
Query Match:            26.1%          Indels:        56
DB:                     3              Gaps:          23

US-10-782-096-2 (1-682) x US-10-089-678-3 (1-3690)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||      |||:::||||||||||||:::|||||||||            ::::::::::::
Db        190 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 246

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              |||:::         ||||||||||||||||   |||            |||::::::|||
Db        247 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 306

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61
              |||||||||::::::|||   |||   :::      |||   :::      |||:::
Db        307 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 366

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
              |||:::   |||                          |||   |||||||||
Db        367 GAGACGTTTATTAGT---------------------TCATCTACGGTTCAAACTGGAATT 405

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                    ||||||   :::|||   ||||||||||||   |||   |||
Db        406 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 465

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
                 |||   ::::::|||   ||||||                      :::|||
Db        466 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 525

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
              :::|||      ||||||:::||||||   :::      |||         ||||||:::
Db        526 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 585

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||      |||:::||||||         :::   :::|||||||         |||:::
Db        586 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 645

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178
                 ||||||   |||:::||||||            |||:::   :::||||||||||||
Db        646 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 705

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
                 ::::::      |||      :::|||   |||                  |||:::
Db        706 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 765
```

```
Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        766 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 825

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        826 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 885

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        886 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 945

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        946 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 1005

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db       1006 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 1065

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db       1066 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 1125

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db       1126 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTTCTTTT 1182

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db       1183 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTTGATATACTCAGTGAAATA 1242

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375
Db       1243 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1302

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
Db       1303 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1362

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
Db       1363 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1419

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
Db       1420 CAGGTTCGATCATTA------------------GGGGCGGATTTAGCTAATTACTACGCA 1461

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
Db       1462 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1521

Qy        455 SerLeu---------ArgTyrGluValProAlaAsn---------------LeuProSer 466
Db       1522 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1581

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
Db       1582 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1635

Qy        487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504
Db       1636 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTTCTAAGAATGCTAGTAGTCCTGCT 1695

Qy        505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524
Db       1696 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1755

Qy        525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541
Db       1756 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1815

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
Db       1816 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1875

Qy        562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577
Db       1876 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1932

Qy        578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597
Db       1933 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1986

Qy        598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617
Db       1987 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGGAT 2037

Qy        618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636
Db       2038 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 2097

Qy        637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653
Db       2098 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 2157

Qy        654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671
Db       2158 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2217

Qy        672 ArgGluGlnGluValAsnAspLeuPhe 680
Db       2218 GCAAAGAAAGCAGTAAATGGCTTGTTT 2244
```

```
RESULT 4
US-09-661-322A-37
; Sequence 37, Application US/09661322A
; Patent No. 6593293
; GENERAL INFORMATION:
;  APPLICANT: Baum, James A.
;  APPLICANT:  Chu, Chih-Rei
;  APPLICANT:  Donovan, William P.
;  APPLICANT:  Gilmer, Amy J.
;  APPLICANT:  Rupar, Mark J.
;  TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Compo
;  TITLE OF INVENTION:  and Methods of Use
;  FILE REFERENCE: MECO201
;  CURRENT APPLICATION NUMBER: US/09/661,322A
;  CURRENT FILING DATE:  2000-09-13
;  NUMBER OF SEQ ID NOS:  63
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
;   LENGTH: 4173
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3687)
US-09-661-322A-37

Alignment Scores:
Pred. No.:              1.25e-86      Length:         4173
Score:                  910.50        Matches:        250
Percent Similarity:     50.4%         Conservative:   103
Best Local Similarity:  35.7%         Mismatches:     280
Query Match:            25.3%         Indels:         67
DB:                     3             Gaps:           22

US-10-782-096-2 (1-682) x US-09-661-322A-37 (1-4173)

Qy          7 LysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyr 26
Db         16 AAAAATGAGAATGAAATTATAAATGCTTTATCGATTCCAGCT---GTATCGAATCATTCC 72

Qy         27 ProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGlu 46
```

```
Db     73 ACACAAATGGATCTATCACCAGATGCTCGTATTGAGGAT-------------------- 111

Qy     47 TrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSer 66

Db    112 ------TCTTTGTGTATAGCCGAGGGGAATAATATCAATCCACTTGTTAGCGCATCAACA 165

Qy     67 ProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThr 86

Db    166 -----------------------------GTCCAAACGGGTATTAACATAGCTGGTAGA 195

Qy     87 IleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeu 106

Db    196 ATACTAGGTGTATTAGGCGTACCGTTTGCTGGACAAATAGCTAGTTTTTATAGTTTTCTT 255

Qy    107 IleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMetValLeuValGlu 125

Db    256 GTTGGTGAATTATGGCCCCGCGGCAGAGATCAGTGGGAAATTTTCCTAGAACATGTCGAA 315

Qy    126 GluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGlu 145

Db    316 CAACTTATAAATCAACAAATAACAGAAAATGCTAGGAATACGGCACTTGCTCGATTACAA 375

Qy    146 GlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLys 165

Db    376 GGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGATTGGCTAGAAAACCGT 435

Qy    166 AsnAspAspAsnArgArgAla---LeuValThrGlnTyrAlaIleValAspAsnPhePhe 184

Db    436 GATGATGCAAGAACGAGAAGTGTTCTTTATACCCAATATATAGCCTTAGAACTTGATTTT 495

Qy    185 GluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyr 204

Db    496 CTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAAGTTCCATTATTAATGGTATAT 555

Qy    205 AlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGln 224

Db    556 GCTCAAGCTGCAAATTTACACCTATTATTATTGAGAGATGCCTCTCTTTTTGGTAGTGAA 615

Qy    225 TrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArg 244

Db    616 TTTGGGCTTACATCGCAGGAAATTCAACGTTATTATGAGCGCCAAGTGGAACAAACGAGA 675

Qy    245 GluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSer 264

Db    676 GATTATTCCGACTATTGCGTAGAATGGTATAATACAGGTCTAAATAGCTTGAGAGGGACA 735

Qy    265 AsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeu 284

Db    736 AATGCCGCAAGTTGGGTGCGTTATAATCAATTCCGTAGAGATCTAACGTTAGGGGTATTA 795

Qy    285 AspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLysThr 304

Db    796 GATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCAATAAATACGAGTGCT 855

Qy    305 GluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGly 324

Db    856 CAGTTAACAAGGGAAGTTTATACAGACGCAATTGGAGCAACAGGGGTA---------AAT 906

Qy    325 GlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsn 344

Db    907 ATGGCAAGTATGAATTGGTATAATAATGCACCTTCGTTTCCGCTATAGAGACTGCG 966

Qy    345 AlaArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThr 364

Db    967 GTTATCCGAAGCCCGCATCTACTTGATTTTCTAGAACAACTTACAATTTTTAGC---ACT 1023

Qy    365 LeuGlyAsnMetSerAspValArgAsnIle------TrpGlyGlyHisThrLeuValGlu 382

Db   1024 TCATCACGATGGAGTGCTACTAGGCATATGACTTACTGGCGGGGGCACACAATTCAATCT 1083

Qy    383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp---SerIleThr 401

Db   1084 CGGCCAATAGGAGGCGGATTAAATACCTCAACGCATGGGTCTACCAATACTTCTATTAAT 1143

Qy    402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421

Db   1144 CCTGTAAGA---TTATCATTCTTCTCTCGAGACGTATATTGGACTGAATCATATGCAGGA 1200

Qy    422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441

Db   1201 GTGCTTCTATGGGGAATT------------TACCTTGAACCTATTCATGGTGTCCCTACT 1248

Qy    442 ValIlePheAsnThrSerAsnIleAsnAsnVal------ProGlySerLeuArgTyrGlu 459

Db   1249 GTTAGATTTAATTTTAGGAACCCTCAGAATACTTTTGAAAGAGGTACTGCTAACTATAGT 1308

Qy    460 ValProAlaAsnLeuProSerGlnThrIleLeu---------SerGluLeuProGlyLys 476

Db   1309 CAACCCTATGAGTCACCTGGGCTTCAATTAAAAGATTCAGAAACTGAATTACCACCAGAA 1368

Qy    477 AspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPhe 496

Db   1369 ACAACAGAACGACCAAATTATGAATCATATAGTCATAGGTTATCTCACATAGGGCTCATT 1428

Qy    497 AspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHis 516

Db   1429 TCACAATCTAGG------------------GTGCATGTACCAGTATATTCTTGGACGCAC 1470

Qy    517 ThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaVal 536

Db   1471 CGTAGTGCAGATCGTACAAATACCATTAGTTCAGATAGCATAACACAAATACCATTGGTA 1530

Qy    537 LysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGlyProThrGlyGly 553

Db   1531 AAATCATTCAACCTTAATTCAGGTACCTCTGTAGTCAGTGGCCCAGGATTTACAGGAGGG 1590

Qy    554 AsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGlnArgGln 573

Db   1591 GATATAATCCGAACTAACGTTAATGGTAGTGTACTAAGTATGGGTCTTAATTTTAATAAT 1650

Qy    574 ThrSer---------TyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAsp 590

Db   1651 ACATCATTACAGCGGTATCGCGTGAGAGTTCGTTATGCTGCTTCTCAAACAATG------ 1704

Qy    591 AlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsn 610

Db   1705 GTCCTGAGGGTAACTGTCGGAGGGAGTACTACTTTTGATCAAGGATTCCCTAGTACTATG 1764

Qy    611 SerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValPro 630

Db   1765 AGTGCA---------AATGAGTCTTTGACATCTCAATCATTTAGATTTGCAGAATTTCCT 1815

Qy    631 ------------------GlyIlePheThrProSerIleAsnProLeuIleArgTyrArg 644

Db   1816 GTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATAAGTAATAATGCAGGT 1875

Qy    645 ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn---Thr 663

Db   1876 AGACAAACGTTT---------CACTTTGATAAAATTGAATTCATTCCAATTACTGCAACC 1926

Qy    664 PheProAsnGln---SerLeuGluLysArgGluGlnGluValAsnAspLeuPheIleAsn 682

Db   1927 TTCGAAGCAGAATACGATTTAGAAAGGGCGCAAGAGGCGGTGAATGCTCTGTTTACTAAT 1986
```

RESULT 5
US-07-876-280-29
; Sequence 29, Application US/07876280
; Patent No. 5262158
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Bagley, Angela L.
; TITLE OF INVENTION: No. 5262158el Bacillus thuringiensis Isolates for
; TITLE OF INVENTION: Controlling Acarides
; NUMBER OF SEQUENCES: 30

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David R. Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/876,280
;      FILING DATE:  19920430
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  M/S 104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3471 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus thuringiensis
;      STRAIN:  kumamotoensis
;      INDIVIDUAL ISOLATE:  PS50C
;    IMMEDIATE SOURCE:
;      CLONE:  E. coli NM522(pMYC2320) NRRL B-18769
US-07-876-280-29

Alignment Scores:
Pred. No.:              8.57e-86       Length:          3471
Score:                  901.50         Matches:         249
Percent Similarity:     49.3%          Conservative:    113
Best Local Similarity:  33.9%          Mismatches:      267
Query Match:            25.1%          Indels:          105
DB:                     2              Gaps:            26

US-10-782-096-2 (1-682) x US-07-876-280-29 (1-3471)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||      |||:::||||||||||||:::::::||||||            ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
                |||:::            ||||||||   |||   :::|||        ::::::|||
Db         61 TCCAGTGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAGATGCGTTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCys---AspSerAsnThrGlnPheIleGlyAspIle 61
                |||||||||::::::|||   |||                 |||   :::      |||:::
Db        121 AATTATAAAGATTATCTGAAAATGTCTGGGGGAGAGAATCCTGAATTATTTGGAAATCCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                   |||:::   |||                                :::   |||||||||
Db        181 GAGACGTTTATTAGT---------------------TCATCCACGATTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                         ||||||   ||||||      |||||||||||   |||||||||
Db        220 GGCATTGTTGGTCGAATACTAGGAGCTTTAGGGGTTCCATTTGCTAGTCAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
                   |||   ::::::|||   ||||||                |||   :::
Db        280 TTCTATAGTTTCATTGTTGGTCAATTATGGCCGTCAAAGAGCGTAGATATATGGGGAGAA 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                :::|||      ||||||||||||:::   :::   |||::::::   |||::::::
Db        340 ATTATGGAACGAGTGGAAGAACTCGTTGATCAAAAAATAGAAAAATATGTAAAAGATAAG 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
                ||||||   ||||||:::||||||            :::   :::||||||      |||:::
Db        400 GCTCTTGCTGAATTAAAAGGGCTAGGAAATGCTTTGGATGTATATCAGCAGTCACTTGAA 459

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThr---GlnTyrAla 178
                   ||||||   |||:::||||||               |||::::::|||:::   |||:::
Db        460 GATTGGCTGGAAAATCGCAATGATGCAAGAACTAGAAGTGTTGTTTCTAATCAATTTATA 519

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
                   :::|||      |||         ::::::|||   |||                |||:::
Db        520 GCTTTAGATCTTAACTTTGTTAGTTCAATTCCATCTTTTGCAGTATCCGGACACGAAGTA 579

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                |||||||||   |||||||||||||||   ||||||||||||:::||||||||||||||||
Db        580 CTATTATTAGCAGTATATGCACAGGCTGTGAACCTACATTTATTGTTATTAAGAGATGCT 639

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                      ||||||   :::|||                  ||||||         |||   |||
Db        640 TCTATTTTTGGAGAAGAGTGGGGATTTACACCAGGTGAAATTTCTAGATTTTATAATCGT 699

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                         |||         ||||||   |||:::|||:::   :::|||      ||||||
Db        700 CAAGTGCAACTTACCGCTGAATATTCAGACTATTGTGTAAAGTGGTATAAAATCGGCTTA 759

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                ::::::         :::   ::::::   |||:::::::::::::::||||||   |||
Db        760 GATAAATTGAAAGGTACCACTTCTAAAAGTTGGCTGAATTATCATCAGTTCCGTAGAGAG 819

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
                |||||||||   ||||||||||||   ||||||||||||||||||            |||
Db        820 ATGACATTACTGGTATTAGATTTGGTGGCGTTATTTCCAAACTATGACACACATATGTAT 879

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
                |||:::            :::|||||||||:::|||||||||||||||:::   |||
Db        880 CCAATCGAAACAACAGCTCAACTTACACGGGATGTGTATACAGATCCGATAGCATTT--- 936

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                   ::::::   |||   |||         ||||||         :::         |||
Db        937 AACATAGTGACAAGTACTGGATTCTGCAACCCTTGGTCAACCCACAGTGGTATTCTTTTT 996

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                   :::|||||||||         |||   |||               |||:::   :::
Db        997 TATGAAGTTGAAAACAACGTAATTCGTCCGCCACACTTGTTTGATATACTCAGCTCAGTA 1056

Qy        359 PheValTyrThrArg------------ThrLeuGlyAsnMetSerAspValArgAsnIle 374
                   :::   |||                  ||||||   |||   :::   :::   |||
Db       1057 GAAATTAATACAAGTAGAGGGGGTATTACGTTAAATAATGATGCATATATA---AACTAC 1113

Qy        375 TrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHis------ 392
                |||   ||||||||||||                  |||   :::|||:::
Db       1114 TGGTCAGGACATACCCTAAAATATCGTAGAACAGCTGATTCGACCGTAACATACACAGCT 1173

Qy        393 AsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSer 412
                |||:::|||:::      |||                              :::
Db       1174 AATTACGGTCGAATCACTTCA-----------------------------GAAAAGAAT 1203

Qy        413 ValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn------ 430
                   |||::::::|||                  :::      |||   ||||||
Db       1204 TCATTTGCACTTGAGGATAGGGATATTTTTGAAATTAATTCAACTGTGGCAAACCTAGCT 1263

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsn 450
                ||||||            ||||||
Db       1264 AATTACTACCAAAAGGCATATGGT------------------------------------ 1287

Qy        451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
                   ||||||||||||      :::   |||               |||   |||
Db       1288 ---GTGCCGGGATCTTGGTTCCATATGGTAAAAAGGGGAACCTCATCAACAACAGCGTAT 1344
```

```
Qy        471 -------------------------------------------------------Ser 471

Db       1345 TTATATTCAAAAACACATACAGCTCTCCAAGGGTGTACACAGGTTTATGAATCAAGTGAT 1404

Qy        472 GluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSer 491
              |||:::|||     |||:::          |||   |||       :::|||||||||||||||
Db       1405 GAAATACCT---CTAGATAGAACTGTACCGGTAGCTGAAAGCTATAGTCATAGATTATCT 1461

Qy        492 TyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThr 511
              :::|||::::::              ::::::   |||            |||
Db       1462 CATATTACCTCCCATTCTTTCTCTAAAAATGGGAGTGCATACTATGGGAGTTTCCCTGTA 1521

Qy        512 PheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThr 531
              |||   |||   ||||||||   |||   ||||||    :::       ||||||||||
Db       1522 TTTGTTTGGACACATACTAGTGCGGATTTAAATAATACAATATATTCAGATAAAATCACT 1581

Qy        532 GlnIleAspAlaValLysGly------------TrpGlyGlyAsnIleGlyPheValIle 547
              ||||||   ||||||||||||              |||||||:::         |||:::
Db       1582 CAAATTCCAGCGGTAAAGGGAGACATGTTATATCTAGGGGGTTCC---------GTAGTA 1632

Qy        548 ProGlyPro------ThrGlyGlyAsnLeuValLysValSerAsp-----------Ser 561
                 ||||||        |||||||||::::::::::|||   ::::::
Db       1633 CAGGGTCCTGGATTTACAGGAGGAGATATATTAAAAAGAACCAATCCTAGCATATTAGGG 1692

Qy        562 TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArg 581
                 ::::::   |||               |||         ||||||:::|||:::|||
Db       1693 ACCTTTGCGGTTACAGTAAATGGTCGTTATCACAA---AGATATCGTGTAAGAATTCGC 1749

Qy        582 TyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHis 601
              ||||||                   |||      |||       :::   |||   :::
Db       1750 TATGCCTCTACAACA---------GATTTTGAATTTACTCTATACCTTGGCGACACAATA 1800

Qy        602 IleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGlu 621
                    :::   |||:::   :::      :::|||             :::|||
Db       1801 GAAAAAAATAGATTTAACAAAACTATGGATAATGGGGCA---------TCTTTAACGTAT 1851

Qy        622 SerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeuIle 641
                       |||:::                        |||         |||
Db       1852 GAAACATTTAAA---------------------TTCGCAAGTTTCATTACTGATTTCCAA 1890

Qy        642 ArgTyrArgThrGln---------------SerPheGlyThrHisAla----------- 652
                       ||||||                   |||   |||         :::
Db       1891 TTCAGAGAAACACAAGATAAAATACTCCTATCCATGGGTGATTTAGCTCCGGTCAAGAA 1950

Qy        653 ------IleAspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSer 668
                    ||||||:::   ||||||||||||::::::   |||:::         |||
Db       1951 GTTTATATAGACCGAATCGAATTCATCCCAGTAGATGAGACATATGAGGCGGAACAAGAT 2010

Qy        669 LeuGluLysArgGluGlnGluValAsnAspLeuPheIleAsn 682
              ||||||      ::::::   ||||||   ||||||   |||
Db       2011 TTAGAAGCGGCGAAGAAAGCAGTGAATGCCTTGTTTACGAAT 2052
```

RESULT 6
US-07-812-180A-1
; Sequence 1, Application US/07812180A
; Patent No. 5366892
; GENERAL INFORMATION:
;   APPLICANT: Foncerrada, Luis R
;   APPLICANT: Payne, Jewel M
;   APPLICANT: Sick, August J
;   TITLE OF INVENTION: No. 5366892el Coleopteran-Active Bacillus
;   TITLE OF INVENTION: thuringiensis Isolate and a No. 5366892el Gene Encoding a
;   TITLE OF INVENTION: Coleopteran-Active Toxin
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Roman Saliwanchik
;   STREET: 2421 N.W. 41st Street, Ste A-1
;   CITY: Gainesville
;   STATE: FL
;   COUNTRY: USA
;   ZIP: 32606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/812,180A
;   FILING DATE: 19920102
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Saliwanchik, Roman
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 904-375-8100
;   TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3471 base pairs
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;   ORGANISM: Bacillus thuringiensis
;   STRAIN: kumamotoensis
;   INDIVIDUAL ISOLATE: PS50C
;   IMMEDIATE SOURCE:
;   LIBRARY: LAMBDAGEM (TM) - 11 LIBRARY OF LUIS
;   LIBRARY: FONCERRADA
;   CLONE: 50C
US-07-812-180A-1

```
Alignment Scores:
Pred. No.:              8.57e-86      Length:         3471
Score:                  901.50        Matches:        249
Percent Similarity:     49.3%         Conservative:   113
Best Local Similarity:  33.9%         Mismatches:     267
Query Match:            25.1%         Indels:         105
DB:                     2             Gaps:           26

US-10-782-096-2 (1-682) x US-07-812-180A-1 (1-3471)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                 |||      |||:::||||||||||||::::::||||||         ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
                 |||:::         |||||||||   |||      :::|||         ::::::|||
Db         61 TCCAGTGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAGATGCGTTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCys---AspSerAsnThrGlnPheIleGlyAspIle 61
              |||||||||::::::|||   |||                |||      |||   |||:::
Db        121 AATTATAAAGATTATCTGAAAATGTCTGGGGGAGAGAATCCTGAATTATTTGGAAATCCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||:::   |||                           :::   |||||||||
Db        181 GAGACGTTTATTAGT--------------------TCATCCACGATTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                    ||||||   ||||||      ||||||||||||   |||||||||
Db        220 GGCATTGTTGGTCGAATACTAGGAGCTTTAGGGGTTCCATTTGCTAGTCAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
                    |||   ::::::|||   ||||||                |||   :::
Db        280 TTCTATAGTTTCATTGTTGGTCAATTATGGCCGTCAAAGAGCGTAGATATATGGGGAGAA 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
              :::|||         ||||||||||||:::   :::   |||::::::   |||::::::
Db        340 ATTATGGAACGAGTGGAAGAACTCGTTGATCAAAAAATAGAAAAATATGTAAAAGATAAG 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||:::||||||         :::   :::||||||      |||:::
Db        400 GCTCTTGCTGAATTAAAAGGGCTAGGAAATGCTTTGGATGTATATCAGCAGTCACTTGAA 459
```

```
Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThr---GlnTyrAla 178
Db        460 GATTGGCTGGAAAATCGCAATGATGCAAGAACTAGAAGTGTTGTTTCTAATCAATTTATA 519

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        520 GCTTTAGATCTTAACTTTGTTAGTTCAATTCCATCTTTTGCAGTATCCGGACACGAAGTA 579

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        580 CTATTATTAGCAGTATATGCACAGGCTGTGAACCTACATTTATTGTTATTAAGAGATGCT 639

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        640 TCTATTTTTGGAGAAGAGTGGGGATTTACACCAGGTGAAATTTCTAGATTTTATAATCGT 699

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        700 CAAGTGCAACTTACCGCTGAATATTCAGACTATTGTGTAAAGTGGTATAAAATCGGCTTA 759

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        760 GATAAATTGAAAGGTACCACTTCTAAAAGTTGGCTGAATTATCATCAGTTCCGTAGAGAG 819

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        820 ATGACATTACTGGTATTAGATTTGGTGGCGTTATTTCCAAACTATGACACACATATGTAT 879

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        880 CCAATCGAAACAACAGCTCAACTTACACGGGATGTGTATACAGATCCGATAGCATTT--- 936

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        937 AACATAGTGACAAGTACTGGATTCTGCAACCCTTGGTCAACCCACAGTGGTATTCTTTTT 996

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        997 TATGAAGTTGAAAACAACGTAATTCGTCCGCCACACTTGTTTGATATACTCAGCTCAGTA 1056

Qy        359 PheValTyrThrArg------------ThrLeuGlyAsnMetSerAspValArgAsnIle 374
Db       1057 GAAATTAATACAAGTAGAGGGGGTATTACGTTAAATAATGATGCATATATA---AACTAC 1113

Qy        375 TrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHis------ 392
Db       1114 TGGTCAGGACATACCCTAAAATATCGTAGAACAGCTGATTCGACCGTAACATACACAGCT 1173

Qy        393 AsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSer 412
Db       1174 AATTACGGTCGAATCACTTCA------------------------------GAAAAGAAT 1203

Qy        413 ValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn------ 430
Db       1204 TCATTTGCACTTGAGGATAGGGATATTTTTGAAATTAATTCAACTGTGGCAAACCTAGCT 1263

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsn 450
Db       1264 AATTACTACCAAAAGGCATATGGT------------------------------------ 1287

Qy        451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
Db       1288 ---GTGCCGGGATCTTGGTTCCATATGGTAAAAAGGGGAACCTCATCAACAACAGCGTAT 1344

Qy        471 ---------------------------------------------------------Ser 471
Db       1345 TTATATTCAAAAACACATACAGCTCTCCAAGGGTGTACACAGGTTTATGAATCAAGTGAT 1404

Qy        472 GluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSer 491
Db       1405 GAAATACCT---CTAGATAGAACTGTACCGGTAGCTGAAAGCTATAGTCATAGATTATCT 1461

Qy        492 TyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThr 511
Db       1462 CATATTACCTCCCATTCTTTCTCTAAAAATGGGAGTGCATACTATGGGAGTTTCCCTGTA 1521

Qy        512 PheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThr 531
Db       1522 TTTGTTTGGACACATACTAGTGCGGATTTAAATAATACAATATATTCAGATAAAATCACT 1581

Qy        532 GlnIleAspAlaValLysGly------------TrpGlyGlyAsnIleGlyPheValIle 547
Db       1582 CAAATTCCAGCGGTAAAGGGAGACATGTTATATCTAGGGGGTTCC---------GTAGTA 1632

Qy        548 ProGlyPro------ThrGlyGlyAsnLeuValLysValSerAsp------------Ser 561
Db       1633 CAGGGTCCTGGATTTACAGGAGGAGATATATTAAAAAGAACCAATCCTAGCATATTAGGG 1692

Qy        562 TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArg 581
Db       1693 ACCTTTGCGGTTACAGTAAATGGGTCGTTATCACAA---AGATATCGTGTAAGAATTCGC 1749

Qy        582 TyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHis 601
Db       1750 TATGCCTCTACAACA---------GATTTTGAATTTACTCTATACCTTGGCGACACAATA 1800

Qy        602 IleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGlu 621
Db       1801 GAAAAAAATAGATTTAACAAAACTATGGATAATGGGGCA---------TCTTTAACGTAT 1851

Qy        622 SerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeuIle 641
Db       1852 GAAACATTTAAA--------------------TTCGCAAGTTTCATTACTGATTTCCAA 1890

Qy        642 ArgTyrArgThrGln---------------SerPheGlyThrHisAla------------ 652
Db       1891 TTCAGAGAAACACAAGATAAAATACTCCTATCCATGGGTGATTTAGCTCCGGTCAAGAA 1950

Qy        653 ------IleAspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSer 668
Db       1951 GTTTATATAGACCGAATCGAATTCATCCCAGTAGATGAGACATATGAGGCGGAACAAGAT 2010

Qy        669 LeuGluLysArgGluGlnGluValAsnAspLeuPheIleAsn 682
Db       2011 TTAGAAGCGGCGAAGAAAGCAGTGAATGCCTTGTTTACGAAT 2052
```

RESULT 7
US-08-315-468-1
; Sequence 1, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Foncerrada, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993

```
;      APPLICATION NUMBER:  07/828,430
;      FILING DATE:  30-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/808,316
;      FILING DATE:  16-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  MA73.C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3471 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus thuringiensis
;      STRAIN:  kumamotoensis
;      INDIVIDUAL ISOLATE:  PS50C
;    IMMEDIATE SOURCE:
;      LIBRARY:  LAMBDAGEM (TM) - 11 LIBRARY OF LUIS
;      LIBRARY:  FONCERRADA
;      CLONE:  50C(a)
US-08-315-468-1

Alignment Scores:
Pred. No.:              8.57e-86      Length:         3471
Score:                  901.50        Matches:        249
Percent Similarity:     49.3%         Conservative:   113
Best Local Similarity:  33.9%         Mismatches:     267
Query Match:            25.1%         Indels:         105
DB:                     2             Gaps:           26

US-10-782-096-2 (1-682) x US-08-315-468-1 (1-3471)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||      |||:::|||||||||||::::::|||||                :::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
                |||:::              ||||||||    |||     :::|||          ::::::|||
Db         61 TCCAGTGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAGATGCGTTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCys---AspSerAsnThrGlnPheIleGlyAspIle 61
                |||||||||::::::|||    |||                |||      :::       |||:::
Db        121 AATTATAAAGATTATCTGAAAATGTCTGGGGGAGAGAATCCTGAATTATTTGGAAATGCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                   |||:::     |||                                  :::     |||||||||
Db        181 GAGACGTTTATTAGT--------------------TCATCCACGATTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                         |||||     ||||||        |||||||||||     |||||||||
Db        220 GGCATTGTTGGTCGAATACTAGGAGCTTTAGGGGTTCCATTTGCTAGTCAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
                         |||     ::::::|||     ||||||                |||      :::
Db        280 TTCTATAGTTTCATTGTTGGTCAATTATGGCCGTCAAAGAGCGTAGATATATGGGGAGAA 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                :::|||          |||||||||||:::      :::      |||::::::      |||::::::
Db        340 ATTATGGAACGAGTGGAAGAACTCGTTGATCAAAAAATAGAAAAATATGTAAAAGATAAG 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
                ||||||     ||||||:::||||||                :::      :::||||||      |||:::
Db        400 GCTCTTGCTGAATTAAAAGGGCTAGGAAATGCTTTGGATGTATATCAGCAGTCACTTGAA 459

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThr---GlnTyrAla 178
                   ||||||     |||:::||||||                |||::::::|||:::     |||:::
Db        460 GATTGGCTGGAAAATCGCAATGATGCAAGAACTAGAAGTGTTGTTTCTAATCAATTTATA 519

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
                   :::|||          |||          ::::::|||    |||                 |||:::
Db        520 GCTTTAGATCTTAACTTTGTTAGTTCAATTCCATCTTTTGCAGTATCCGGACACGAAGTA 579

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                |||||||||     ||||||||||||||     ||||||||||||:::||||||||||||||
Db        580 CTATTATTAGCAGTATATGCACAGGCTGTGAACCTACATTTATTGTTATTAAGAGATGCT 639

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                         ||||||     :::|||                   ||||||          |||    |||
Db        640 TCTATTTTTGGAGAAGAGTGGGGATTTACACCAGGTGAAATTTCTAGATTTTATAATCGT 699

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                            |||          ||||||     |||:::|||:::    :::|||          ||||||
Db        700 CAAGTGCAACTTACCGCTGAATATTCAGACTATTGTGTAAAGTGGTATAAAATCGGCTTA 759

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                ::::::                  :::      ::::::      |||::::::::::::::::|||||||    :::
Db        760 GATAAATTGAAAGGTACCACTTCTAAAAGTTGGCTGAATTATCATCAGTTCCGTAGAGAG 819

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
                |||||||||     ||||||||||||||     |||||||||||||||||                 |||
Db        820 ATGACATTACTGGTATTAGATTTGGTGGCGTTATTTCCAAACTATGACACACATATGTAT 879

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
                |||:::                   :::|||||||||:::||||||||||||||||:::      |||
Db        880 CCAATCGAAACAACAGCTCAACTTACACGGGATGTGTATACAGATCCGATAGCATTT--- 936

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                   ::::::    |||     |||               ||||||               :::             |||
Db        937 AACATAGTGACAAGTACTGGATTCTGCAACCCTTGGTCAACCCACAGTGGTATTCTTTTT 996

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                      :::|||||||||          |||    |||                         |||:::     :::
Db        997 TATGAAGTTGAAAACAACGTAATTCGTCCGCCACACTTGTTTGATATACTCAGCTCAGTA 1056

Qy        359 PheValTyrThrArg-----------ThrLeuGlyAsnMetSerAspValArgAsnIle 374
                 :::      |||                    ||||||      |||      :::     :::    |||
Db       1057 GAAATTAATACAAGTAGAGGGGGTATTACGTTAAATAATGATGCATATATA---AACTAC 1113

Qy        375 TrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHis------ 392
                |||     |||||||||||||                         |||     :::|||:::
Db       1114 TGGTCAGGACATACCCTAAAATATCGTAGAACAGCTGATTCGACCGTAACATACACAGCT 1173

Qy        393 AsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSer 412
                |||:::|||:::          |||                                             :::
Db       1174 AATTACGGTCGAATCACTTCA-----------------------------GAAAAGAAT 1203

Qy        413 ValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn------ 430
                  |||::::::|||                  :::             |||    ||||||
Db       1204 TCATTTGCACTTGAGGATAGGGATATTTTTGAAATTAATTCAACTGTGGCAAACCTAGCT 1263

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsn 450
                ||||||                |||||
Db       1264 AATTACTACCAAAAGGCATATGGT----------------------------------- 1287

Qy        451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
                   ||||||||||||          :::     |||                   |||     |||
Db       1288 ---GTGCCGGGATCTTGGTTCCATATGGTAAAAAGGGGAACCTCATCAACAACAGCGTAT 1344

Qy        471 ---------------------------------------------------------Ser 471

Db       1345 TTATATTCAAAAACACATACAGCTCTCCAAGGGTGTACACAGGTTTATGAATCAAGTGAT 1404

Qy        472 GluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSer 491
                |||:::|||        |||:::               |||     |||         :::|||||||||||||
Db       1405 GAAATACCT---CTAGATAGAACTGTACCGGTAGCTGAAAGCTATAGTCATAGATTATCT 1461

Qy        492 TyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThr 511
```

```
                 :::|||::::::             ::::::   |||          |||
Db       1462 CATATTACCTCCCATTCTTTCTCTAAAAATGGGAGTGCATACTATGGGAGTTTCCCTGTA 1521

Qy        512 PheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThr 531
              |||  |||  |||||||||  |||  ||||||  :::    |||||||||||
Db       1522 TTTGTTTGGACACATACTAGTGCGGATTTAAATAATACAATATATTCAGATAAAATCACT 1581

Qy        532 GlnIleAspAlaValLysGly------------TrpGlyGlyAsnIleGlyPheValIle 547
              ||||||  ||||||||||||              ||||||:::          |||:::
Db       1582 CAAATTCCAGCGGTAAAGGGAGACATGTTATATCTAGGGGGTTCC---------GTAGTA 1632

Qy        548 ProGlyPro------ThrGlyGlyAsnLeuValLysValSerAsp------------Ser 561
                 ||||||       ||||||||:::::::::|||    ::::::
Db       1633 CAGGGTCCTGGATTTACAGGAGGAGATATATTAAAAAGAACCAATCCTAGCATATTAGGG 1692

Qy        562 TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArg 581
                    ::::::  |||                |||     ||||||:::|||:::|||
Db       1693 ACCTTTGCGGTTACAGTAAATGGGTCGTTATCACAA---AGATATCGTGTAAGAATTCGC 1749

Qy        582 TyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHis 601
              ||||||                   |||    |||       :::   |||    :::
Db       1750 TATGCCTCTACAACA---------GATTTTGAATTTACTCTATACCTTGGCGACACAATA 1800

Qy        602 IleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGlu 621
                    :::  ||||:::   :::        :::|||             :::|||
Db       1801 GAAAAAAATAGATTTAACAAAACTATGGATAATGGGGCA---------TCTTTAACGTAT 1851

Qy        622 SerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeuIle 641
                    |||:::                     |||       |||
Db       1852 GAAACATTTAAA---------------------TTCGCAAGTTTCATTACTGATTTCCAA 1890

Qy        642 ArgTyrArgThrGln---------------SerPheGlyThrHisAla------------ 652
                       ||||||                |||   |||      :::
Db       1891 TTCAGAGAAACACAAGATAAAATACTCCTATCCATGGGTGATTTAGCTCCGGTCAAGAA 1950

Qy        653 ------IleAspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSer 668
                    ||||||:::    ||||||||||||::::::     |||:::         |||
Db       1951 GTTTATATAGACCGAATCGAATTCATCCCAGTAGATGAGACATATGAGGCGGAACAAGAT 2010

Qy        669 LeuGluLysArgGluGlnGluValAsnAspLeuPheIleAsn 682
              ||||||          ::::::   ||||||   ||||||   |||
Db       2011 TTAGAAGCGGCGAAGAAAGCAGTGAATGCCTTGTTTACGAAT 2052
```

RESULT 8
US-07-941-650A-1
; Sequence 1, Application US/07941650A
; Patent No. 6294184
; GENERAL INFORMATION:
; APPLICANT: Uyeda, Kendrick A.
; APPLICANT: Bradfisch, Gregory A.
; TITLE OF INVENTION: Process for Controlling Lepidopteran Pests
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: David R. Saliwanchik
;   STREET: 2421 N.W. 41st Street, Suite A-1
;   CITY: Gainesville
;   STATE: FL
;   COUNTRY: USA
;   ZIP: 32606
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/941,650A
;   FILING DATE: 19920908
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/758,020
;   FILING DATE: 12-SEP-1991
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/658,935
;   FILING DATE: 21-FEB-1991
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/642,112
;   FILING DATE: 16-JAN-1991
; ATTORNEY/AGENT INFORMATION:
;   NAME: Saliwanchik, David R.
;   REGISTRATION NUMBER: 31,794
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 904-375-8100
;   TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3471 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
;   ORGANISM: Bacillus thuringiensis
;   STRAIN: kumamotoensis
;   INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
;   CLONE: E. coli NM522(pMYC1638), NRRL B-18751
US-07-941-650A-1

```
Alignment Scores:
Pred. No.:              8.57e-86      Length:        3471
Score:                  901.50        Matches:       249
Percent Similarity:     49.3%         Conservative:  113
Best Local Similarity:  33.9%         Mismatches:    267
Query Match:            25.1%         Indels:        105
DB:                     3             Gaps:          26

US-10-782-096-2 (1-682) x US-07-941-650A-1 (1-3471)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
               |||     |||:::|||||||||||||::::::||||||            ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              |||:::            |||||||||   |||    :::|||          ::::::|||
Db         61 TCCAGTGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAGATGCGTTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCys---AspSerAsnThrGlnPheIleGlyAspIle 61
              ||||||||::::::|||   |||                |||     :::     |||:::
Db        121 AATTATAAAGATTATCTGAAAATGTCTGGGGGAGAGAATCCTGAATTATTTGGAAATCCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||:::   |||                                :::   |||||||||
Db        181 GAGACGTTTATTAGT---------------------TCATCCACGATTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                    ||||||   ||||||     ||||||||||||   |||||||||
Db        220 GGCATTGTTGGTCGAATACTAGGAGCTTTAGGGGTTCCATTTGCTAGTCAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
                    |||    ::::::|||   ||||||                |||   :::
Db        280 TTCTATAGTTTCATTGTTGGTCAATTATGGCCGTCAAAGAGCGTAGATATATGGGGAGAA 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
              :::|||      ||||||||||||:::   :::   |||::::::   |||::::::
Db        340 ATTATGGAACGAGTGGAAGAACTCGTTGATCAAAAAATAGAAAAATATGTAAAAGATAAG 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||:::|||||            :::     :::||||||   |||:::
Db        400 GCTCTTGCTGAATTAAAAGGGCTAGGAAATGCTTTGGATGTATATCAGCAGTCACTTGAA 459

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThr---GlnTyrAla 178
                 ||||||   |||:::||||||            |||::::::|||:::   |||:::
Db        460 GATTGGCTGGAAAATCGCAATGATGCAAGAACTAGAAGTGTTGTTTCTAATCAATTTATA 519
```

```
Qy    179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db    520 GCTTTAGATCTTAACTTTGTTAGTTCAATTCCATCTTTTGCAGTATCCGGACACGAAGTA 579

Qy    199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db    580 CTATTATTAGCAGTATATGCACAGGCTGTGAACCTACATTTATTGTTATTAAGAGATGCT 639

Qy    219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db    640 TCTATTTTTGGAGAAGAGTGGGGATTTACACCAGGTGAAATTTCTAGATTTTATAATCGT 699

Qy    239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db    700 CAAGTGCAACTTACCGCTGAATATTCAGACTATTGTGTAAAGTGGTATAAAATCGGCTTA 759

Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db    760 GATAAATTGAAAGGTACCACTTCTAAAAGTTGGCTGAATTATCATCAGTTCCGTAGAGAG 819

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db    820 ATGACATTACTGGTATTAGATTTGGTGGCGTTATTTCCAAACTATGACACACATATGTAT 879

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db    880 CCAATCGAAACAACAGCTCAACTTACACGGGATGTGTATACAGATCCGATAGCATTT--- 936

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db    937 AACATAGTGACAAGTACTGGATTCTGCAACCCTTGGTCAACCCACAGTGGTATTCTTTTT 996

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db    997 TATGAAGTTGAAAACAACGTAATTCGTCCGCCACACTTGTTTGATATACTCAGCTCAGTA 1056

Qy    359 PheValTyrThrArg------------ThrLeuGlyAsnMetSerAspValArgAsnIle 374
Db   1057 GAAATTAATACAAGTAGAGGGGGTATTACGTTAAATAATGATGCATATATA---AACTAC 1113

Qy    375 TrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHis------ 392
Db   1114 TGGTCAGGACATACCCTAAAATATCGTAGAACAGCTGATTCGACCGTAACATACACACCT 1173

Qy    393 AsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSer 412
Db   1174 AATTACGGTCGAATCACTTCA-----------------------------GAAAAGAAT 1203

Qy    413 ValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn------ 430
Db   1204 TCATTTGCACTTGAGGATAGGGATATTTTTGAAATTAATTCAACTGTGGCAAACCTAGCT 1263

Qy    431 AsnTyrIleThrSerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsn 450
Db   1264 AATTACTACCAAAAGGCATATGGT------------------------------------ 1287

Qy    451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
Db   1288 ---GTGCCGGGATCTTGGTTCCATATGGTAAAAAGGGGAACCTCATCAACAACAGCGTAT 1344

Qy    471 ---------------------------------------------------------Ser 471
Db   1345 TTATATTCAAAAACACATACAGCTCTCCAAGGGTGTACACAGGTTTATGAATCAAGTGAT 1404

Qy    472 GluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSer 491
Db   1405 GAAATACCT---CTAGATAGAACTGTACCGGTAGCTGAAAGCTATAGTCATAGATTATCT 1461

Qy    492 TyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThr 511
Db   1462 CATATTACCTCCCATTCTTTCTCTAAAAATGGGAGTGCATACTATGGGAGTTTCCCTGTA 1521

Qy    512 PheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThr 531
Db   1522 TTTGTTTGGACACATACTAGTGCGGATTTAAATAATACAATATATTCAGATAAAATCACT 1581

Qy    532 GlnIleAspAlaValLysGly------------TrpGlyGlyAsnIleGlyPheValIle 547
Db   1582 CAAATTCCAGCGGTAAAGGGAGACATGTTATATCTAGGGGGTTCC---------GTAGTA 1632

Qy    548 ProGlyPro------ThrGlyGlyAsnLeuValLysValSerAsp-----------Ser 561
Db   1633 CAGGGTCCTGGATTTACAGGAGGAGATATATTAAAAAGAACCAATCCTAGCATATTAGGG 1692

Qy    562 TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArg 581
Db   1693 ACCTTTGCGGTTACAGTAAATGGGTCGTTATCACAA---AGATATCGTGTAAGAATTCGC 1749

Qy    582 TyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHis 601
Db   1750 TATGCCTCTACAACA---------GATTTTGAATTTACTCTATACCTTGGCGACACAATA 1800

Qy    602 IleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGlu 621
Db   1801 GAAAAAAATAGATTTAACAAAACTATGGATAATGGGGCA---------TCTTTAACGTAT 1851

Qy    622 SerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeuIle 641
Db   1852 GAAACATTTAAA--------------------TTCGCAAGTTTCATTACTGATTTCCAA 1890

Qy    642 ArgTyrArgThrGln---------------SerPheGlyThrHisAla------------ 652
Db   1891 TTCAGAGAAACACAAGATAAAATACTCCTATCCATGGGTGATTTAGCTCCGGTCAAGAA 1950

Qy    653 ------IleAspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSer 668
Db   1951 GTTTATATAGACCGAATCGAATTCATCCCAGTAGATGAGACATATGAGGCGGAACAAGAT 2010

Qy    669 LeuGluLysArgGluGlnGluValAsnAspLeuPheIleAsn 682
Db   2011 TTAGAAGCGGCGAAGAAAGCAGTGAATGCCTTGTTTACGAAT 2052
```

```
RESULT 9
US-08-448-170-7
; Sequence 7, Application US/08448170
; Patent No. 5723758
;  GENERAL INFORMATION:
;    APPLICANT:  Payne, Jewel
;    APPLICANT:  Cummings, David A.
;    APPLICANT:  Cannon, Raymond J.C.
;    APPLICANT:  Narva, Kenneth E.
;    APPLICANT:  Stelman, Steve
;    TITLE OF INVENTION:  No. 5723758el Bacillus thuringiensis Isolate Denoted
;    TITLE OF INVENTION:  B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
;    TITLE OF INVENTION:  Encoding Lepidopteran-Active Toxins
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David R. Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,170
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/069,902
;      FILING DATE:  01-JUNE-1993
;      CLASSIFICATION:  424
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/759,247
;      FILING DATE:  13-SEPT-1991
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  M/S 102D.C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (904) 375-8100
;      TELEFAX:  (904) 372-5800
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3684 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-448-170-7

Alignment Scores:
Pred. No.:              2.86e-85      Length:        3684
Score:                  897.00        Matches:       252
Percent Similarity:     50.7%         Conservative:  104
Best Local Similarity:  35.9%         Mismatches:    274
Query Match:            24.9%         Indels:        72
DB:                     2             Gaps:          24

US-10-782-096-2 (1-682) x US-08-448-170-7 (1-3684)

Qy      7 LysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyr 26
Db     16 AAAAATGAGAATGAAATTATAAATGCTTTATCGATTCCAGCT---GTATCGAATCATTCC 72

Qy     27 ProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGlu 46
Db     73 GCACAAATGAATCTATCAACCGATGCTCGTATT--------------------GAGGAT 111

Qy     47 TrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSer 66
Db    112 AGCTTGTGTATAGCCGAGGGGAAC---------------AATATCGATCCATTTGTTAGC 156

Qy     67 ProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThr 86
Db    157 ---------------------GCATCAACAGTCCAAACGGGTATTAACATAGCTGGTAGA 195

Qy     87 IleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeu 106
Db    196 ATACTAGGTGTATTAGGCGTACCGTTTGCTGGACAAATAGCTAGTTTTATAGTTTTCTT 255

Qy    107 IleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMetValLeuValGlu 125
Db    256 GTTGGTGAATTATGGCCCCGCGGCAGAGATCCTTGGGAAATTTTCCTAGAACATGTCGAA 315

Qy    126 GluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGlu 145
Db    316 CAACTTATAAGACAACAAGTAACAGAAAATACTAGGGATACGGCTCTTGCTCGATTACAA 375

Qy    146 GlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLys 165
Db    376 GGTTTAGGAAATTCCTTTAGAGCCTATCAACAGTCACTTGAAGATTGGCTAGAAAACCGT 435

Qy    166 AsnAspAspAsnArgArgAla---LeuValThrGlnTyrAlaIleValAspAsnPhePhe 184
Db    436 GATGATGCAAGAACGAGAAGTGTTCTTTATACCCAATATATAGCCTTAGAACTTGATTTT 495

Qy    185 GluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyr 204
Db    496 CTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAAGTTCCATTATTAATGGTATAT 555

Qy    205 AlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGln 224
Db    556 GCTCAAGCTGCAAATTTACACCTATTATTATTGAGAGATGCCTCTCTTTTTGGTAGTGAA 615

Qy    225 TrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArg 244
Db    616 TTTGGGCTTACATCCCAAGAAATTCAACGTTATTATGAGCGCCAAGTGGAAAAAACGAGA 675

Qy    245 GluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSer 264
Db    676 GAATATTCTGATTATTGCGCAAGATGGTATAATACGGGTTTAAATAATTTGAGAGGGACA 735

Qy    265 AsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeu 284
Db    736 AATGCTGAAAGTTGGTTGCGATATAATCAATTCCGTAGAGACTTAACGCTAGGAGTATTA 795

Qy    285 AspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLysThr 304
Db    796 GATCTAGTGGCACTATTCCCAAGCTATGACACGCGTGTTTATCCAATGAATACCAGTGCT 855

Qy    305 GluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGly 324
Db    856 CAATTAACAAGAGAAATTTATACAGATCCAATTGGGAGAACAAATGCA---CCTTCAGGA 912

Qy    325 GlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsn 344
Db    913 TTTGCAAGTACGAATTGGTTTAATAATAATGCACCATCGTTTTCTGCCATAGAGGCTGCC 972

Qy    345 AlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThr 364
Db    973 GTTATTAGGCCTCCGCATCTACTTGATTTTCCAGAACAGCTTACAATTTTCAGC---GTA 1029

Qy    365 LeuGlyAsnMetSerAspValArg------AsnIleTrpGlyGlyHisThrLeuValGlu 382
Db   1030 TTAAGTCGATGGAGTAATACTCAATATATGAATTACTGGGTGGGACATAGACTTGAATCG 1089

Qy    383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp---SerIleThr 401
Db   1090 CGAACAATAAGGGGGTCATTAAGTACCTCGACACACGGAAATACCAATACTTCTATTAAT 1149

Qy    402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
Db   1150 CCTGTAACA---TTACAGTTCACATCTCGAGACGTTTATAGAACAGAATCATTTGCAGGG 1206

Qy    422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
Db   1207 ATAAATATACTTCTAACTACTCCTGTGAAT------------------GGAGTACCTTGG 1248

Qy    442 ValIlePheAsnThrSerAsn---IleAsnAsnValProGlySerLeuArgTyrGluVal 460
Db   1249 GCTAGATTTAATTGGAGAAATCCCCTGAATTCTCTTAGAGGTAGCCTTCTCTATACTATA 1308

Qy    461 ProAlaAsn------------------LeuProSerGlnThrIleLeuSerGluLeuPro 474
Db   1309 GGGTATACTGGAGTGGGGACACAACTATTTGATTCAGAAACT---------GAATTACCA 1359

Qy    475 GlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSer 494
Db   1360 CCAGAAACAACAGAACGACCAAATTATGAATCTTACAGTCATAGATTATCTAATATA--- 1416

Qy    495 AsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrp 514
Db   1417 ---------------AGACTAATATCAGGAAACACTTTGAGAGCACCAGTATATTCTTGG 1461

Qy    515 AlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAsp 534
Db   1462 ACGCACCGTAGTGCAGATCGTACAAATACCATTAGTTCAGATAGCATAACACAAATACCA 1521

Qy    535 AlaValLysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGlyProThr 551
Db   1522 TTGGTAAAATCATTCAACCTTAATTCAGGTACCTCTGTAGTCAGTGGCCCAGGATTTACA 1581

Qy    552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
Db   1582 GGAGGGGATATAATCCGAACTAACGTTAATGGTAGTGTACTAAGTATGGGTCTTAATTTT 1641
```

```
Qy        572 ArgGlnThrSer---------TyrArgIleArgLeuArgTyrAlaCysLeuValThrHis 588
Db       1642 AATAATACATCATTACAGCGGTATCGCGTGAGAGTTCGTTATGCTGCTTCTCAAACAATG 1701

Qy        589 GlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCys 608
Db       1702 ------GTCCTGAGGGTAACTGTCGGAGGGAGTACTACTTTTGATCAAGGATTCCCTAGT 1755

Qy        609 SerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAsp 628
Db       1756 ACTATGAGTGCA---------AATGAGTCTTTGACATCTCAATCATTTAGATTTGCAGAA 1806

Qy        629 ValPro------------------GlyIlePheThrProSerIleAsnProLeuIleArg 642
Db       1807 TTTCCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATAAGTAATAAT 1866

Qy        643 TyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db       1867 GCAGGTAGACAAACGTTT---------CACTTTGATAAAATTGAATTCATTCCAATTACT 1917

Qy        663 ---ThrPheProAsnGln---SerLeuGluLysArgGluGlnGluValAsnAspLeuPhe 680
Db       1918 GCAACCTTCGAAGCAGAATATGATTTAGAAAGAGCGCAAGAGGCGGTGAATGCTCTGTTT 1977

Qy        681 IleAsn 682
Db       1978 ACTAAT 1983
```

RESULT 10
US-08-961-803-5
; Sequence 5, Application US/08961803
; Patent No. 6150589
; GENERAL INFORMATION:
;   APPLICANT: Payne, Jewel
;   APPLICANT: Cummings, David A.
;   APPLICANT: Cannon, Raymond J.C.
;   APPLICANT: Narva, Kenneth E.
;   APPLICANT: Stelman, Steve
;   TITLE OF INVENTION: No. 6150589el Bacillus thuringiensis Isolate Denoted
;   TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
;   TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Jay M. Sanders
;   STREET: 2421 N.W. 41st Street, Suite A-1
;   CITY: Gainesville
;   STATE: Florida
;   COUNTRY: USA
;   ZIP: 32606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,803
;   FILING DATE: 31-OCT-1997
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/069,902
;   FILING DATE: 01-JUNE-1993
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/759,247
;   FILING DATE: 13-SEPT-1991
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/448,170
;   FILING DATE: 23-MAY-1995
;   CLASSIFICATION: 800
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sanders, Jay M.
;   REGISTRATION NUMBER: 39,355
;   REFERENCE/DOCKET NUMBER: M/S 102DCD1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (352) 375-8100
;   TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3684 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-961-803-5

Alignment Scores:

Pred. No.:	2.86e-85	Length:	3684
Score:	897.00	Matches:	252
Percent Similarity:	50.7%	Conservative:	104
Best Local Similarity:	35.9%	Mismatches:	274
Query Match:	24.9%	Indels:	72
DB:	3	Gaps:	24

US-10-782-096-2 (1-682) x US-08-961-803-5 (1-3684)

```
Qy          7 LysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyr 26
Db         16 AAAAATGAGAATGAAATTATAAATGCTTTATCGATTCCAGCT---GTATCGAATCATTCC 72

Qy         27 ProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGlu 46
Db         73 GCACAAATGAATCTATCAACCGATGCTCGTATT--------------------GAGGAT 111

Qy         47 TrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSer 66
Db        112 AGCTTGTGTATAGCCGAGGGAAC---------------AATATCGATCCATTTGTTAGC 156

Qy         67 ProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThr 86
Db        157 --------------------GCATCAACAGTCCAAACGGGTATTAACATAGCTGGTAGA 195

Qy         87 IleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeu 106
Db        196 ATACTAGGTGTATTAGGCGTACCGTTTGCTGGACAAATAGCTAGTTTTTATAGTTTTCTT 255

Qy        107 IleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMetValLeuValGlu 125
Db        256 GTTGGTGAATTATGGCCCCGCGGCAGAGATCCTTGGGAAATTTTCCTAGAACATGTCGAA 315

Qy        126 GluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGlu 145
Db        316 CAACTTATAAGACAACAAGTAACAGAAAATACTAGGGATACGGCTCTTGCTCGATTACAA 375

Qy        146 GlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLys 165
Db        376 GGTTTAGGAAATTCCTTTAGAGCCTATCAACAGTCACTTGAAGATTGGCTAGAAAACCGT 435

Qy        166 AsnAspAspAsnArgArgAla---LeuValThrGlnTyrAlaIleValAspAsnPhePhe 184
Db        436 GATGATGCAAGAACGAGAAGTGTTCTTTATACCCAATATATAGCCTTAGAACTTGATTTT 495

Qy        185 GluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyr 204
Db        496 CTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAAGTTCCATTATTAATGGTATAT 555

Qy        205 AlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGln 224
Db        556 GCTCAAGCTGCAAATTTACACCTATTATTATTGAGAGATGCCTCTCTTTTTGGTAGTGAA 615

Qy        225 TrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArg 244
Db        616 TTTGGGCTTACATCCCAAGAAATTCAACGTTATTATGAGCGCCAAGTGGAAAAAACGAGA 675

Qy        245 GluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSer 264
```

```
Db        676 GAATATTCTGATTATTGCGCAAGATGGTATAATACGGGTTTAAATAATTTGAGAGGGACA 735

Qy        265 AsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeu 284
              ||||||:::   |||:::   :::|||:::|||||   |||:::||||||   ||||||
Db        736 AATGCTGAAAGTTGGTTGCGATATAATCAATTCCGTAGAGACTTAACGCTAGGAGTATTA 795

Qy        285 AspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLysThr 304
              ||||||      |||||||||:::|||||   |||   ||||||:::
Db        796 GATCTAGTGGCACTATTCCCAAGCTATGACACGCGTGTTTATCCAATGAATACCAGTGCT 855

Qy        305 GluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGly 324
              :::|||||||||||||:::|||||||||||||:::|||   |||            ||||||
Db        856 CAATTAACAAGAGAAATTTATACAGATCCAATTGGGAGAACAAATGCA---CCTTCAGGA 912

Qy        325 GlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsn 344
                 :::      |||:::|||   |||      :::|||:::|||:::|||
Db        913 TTTGCAAGTACGAATTGGTTTAATAATAATGCACCATCGTTTTCTGCCATAGAGGCTGCC 972

Qy        345 AlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThr 364
                    |||   |||            :::      ::::::   :::::::::
Db        973 GTTATTAGGCCTCCGCATCTACTTGATTTTCCAGAACAGCTTACAATTTTCAGC---GTA 1029

Qy        365 LeuGlyAsnMetSerAspValArg------AsnIleTrpGlyGlyHisThrLeuValGlu 382
              |||      |||:::   :::         |||   |||   ||||||   |||
Db       1030 TTAAGTCGATGGAGTAATACTCAATATATGAATTACTGGGTGGGACATAGACTTGAATCG 1089

Qy        383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp---SerIleThr 401
                          ||||||      |||         |||   |||:::   ||||||
Db       1090 CGAACAATAAGGGGGTCATTAAGTACCTCGACACACGGAAATACCAATACTTCTATTAAT 1149

Qy        402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
              |||:::         |||   :::      |||:::      ||||||   |||
Db       1150 CCTGTAACA---TTACAGTTCACATCTCGAGACGTTTATAGAACAGAATCATTTGCAGGG 1206

Qy        422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
              |||   :::   |||         |||                  ||||||
Db       1207 ATAAATATACTTCTAACTACTCCTGTGAAT------------------GGAGTACCTTGG 1248

Qy        442 ValIlePheAsnThrSerAsn---IleAsnAsnValProGlySerLeuArgTyrGluVal 460
                 ||||||   |||   :::|||::::::   ||||||||   |||   :::
Db       1249 GCTAGATTTAATTGGAGAAATCCCCTGAATTCTCTTAGAGGTAGCCTTCTCTATACTATA 1308

Qy        461 ProAlaAsn------------------LeuProSerGlnThrIleLeuSerGluLeuPro 474
                                            |||:::|||      |||||||||
Db       1309 GGGTATACTGGAGTGGGGACACAACTATTTGATTCAGAAACT---------GAATTACCA 1359

Qy        475 GlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSer 494
                 :::      |||||||||      :::||||||||||||||   |||
Db       1360 CCAGAAACAACAGAACGACCAAATTATGAATCTTACAGTCATAGATTATCTAATATA--- 1416

Qy        495 AsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrp 514
                          |||   ||||||      :::            :::   |||
Db       1417 ---------------AGACTAATATCAGGAAACACTTTGAGAGCACCAGTATATTCTTGG 1461

Qy        515 AlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAsp 534
                 |||   |||   ||||||   |||   :::      |||   ||||||||||||
Db       1462 ACGCACCGTAGTGCAGATCGTACAAATACCATTAGTTCAGATAGCATAACACAAATACCA 1521

Qy        535 AlaValLysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGlyProThr 551
                 ||||||   :::      |||   |||      :::         ||||||   |||
Db       1522 TTGGTAAAATCATTCAACCTTAATTCAGGTACCTCTGTAGTCAGTGGCCCAGGATTTACA 1581

Qy        552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
              ||||||:::::::::::   :::   :::   |||      :::
Db       1582 GGAGGGGATATAATCCGAACTAACGTTAATGGTAGTGTACTAAGTATGGGTCTTAATTTT 1641

Qy        572 ArgGlnThrSer---------TyrArgIleArgLeuArgTyrAlaCysLeuValThrHis 588
                 ||||||            |||||:::|||:::|||||||||         |||
Db       1642 AATAATACATCATTACAGCGGTATCGCGTGAGAGTTCGTTATGCTGCTTCTCAAACAATG 1701

Qy        589 GlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCys 608
                    :::   |||         ||||||:::                  |||
Db       1702 ------GTCCTGAGGGTAACTGTCGGAGGGAGTACTACTTTTGATCAAGGATTCCCTAGT 1755

Qy        609 SerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAsp 628
              :::   |||:::         :::   :::|||               ||||||:::   :::
Db       1756 ACTATGAGTGCA---------AATGAGTCTTTGACATCTCAATCATTTAGATTTGCAGAA 1806

Qy        629 ValPro------------------GlyIlePheThrProSerIleAsnProLeuIleArg 642
                 |||                  |||      |||      |||:::
Db       1807 TTTCCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATAAGTAATAAT 1866

Qy        643 TyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
                    |||:::|||                  ||||||   ||||||||||||:::
Db       1867 GCAGGTAGACAAACGTTT---------CACTTTGATAAAATTGAATTCATTCCAATTACT 1917

Qy        663 ---ThrPheProAsnGln---SerLeuGluLysArgGluGlnGluValAsnAspLeuPhe 680
                 ||||||   :::      ||||||:::   ::::::   ||||||   ||||||
Db       1918 GCAACCTTCGAAGCAGAATATGATTTAGAAAGAGCGCAAGAGGCGGTGAATGCTCTGTTT 1977

Qy        681 IleAsn 682
                 |||
Db       1978 ACTAAT 1983
```

RESULT 11
US-09-661-322A-62
; Sequence 62, Application US/09661322A
; Patent No. 6593293
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Compo
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: MECO201
; CURRENT APPLICATION NUMBER: US/09/661,322A
; CURRENT FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 62
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-661-322A-62

Alignment Scores:

Pred. No.:	6e-85	Length:	3684
Score:	894.00	Matches:	252
Percent Similarity:	50.6%	Conservative:	103
Best Local Similarity:	35.9%	Mismatches:	275
Query Match:	24.9%	Indels:	72
DB:	3	Gaps:	24

US-10-782-096-2 (1-682) x US-09-661-322A-62 (1-3684)

```
Qy          7 LysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyr 26
              |||||||||   |||:::::::::||||||   |||   :::   :::||||||
Db         16 AAAAATGAGAATGAAATTATAAATGCTTTATCGATTCCAGCT---GTATCGAATCATTCC 72

Qy         27 ProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGlu 46
                 :::      |||:::   |||   ::::::                     ::::::
Db         73 GCACAAATGAATCTATCAACCGATGCTCGTATT---------------------GAGGAT 111

Qy         47 TrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSer 66
                 |||   :::   |||                  :::|||   :::   |||
Db        112 AGCTTGTGTATAGCCGAGGGGAAC---------------AATATCGATCCATTTGTTAGC 156

Qy         67 ProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThr 86
                                         |||   ||||||||||||      |||
Db        157 ---------------------GCATCAACAGTCCAAACGGGTATTAACATAGCTGGTAGA 195

Qy         87 IleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeu 106
              ||||||   ||||||||||||   |||   |||               |||   |||
Db        196 ATACTAGGTGTATTAGGCGTACCGTTTGCTGGACAAATAGCTAGTTTTTATAGTTTTCTT 255
```

```
Qy        107 IleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMetValLeuValGlu 125
Db        256 GTTGGTGAATTATGGCCCCGCGGCAGAGATCCTTGGGAAATTTCCTAGAACATGTCGAA 315

Qy        126 GluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGlu 145
Db        316 CATCTTATAAGACAACAAGTAACAGAAAATACTAGGGATACGGCTCTTGCTCGATTACAA 375

Qy        146 GlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLys 165
Db        376 GGTTTAGGAAATTCCTTTAGAGCCTATCAACAGTCACTTGAAGATTGGCTAGAAAACCGT 435

Qy        166 AsnAspAspAsnArgArgAla---LeuValThrGlnTyrAlaIleValAspAsnPhePhe 184
Db        436 GATGATGCAAGAACGAGAAGTGTTCTTTATACCCAATATATAGCCTTAGAACTTGATTTT 495

Qy        185 GluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyr 204
Db        496 CTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAAGTTCCATTATTAATGGTATAT 555

Qy        205 AlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGln 224
Db        556 GCTCAAGCTGCAAATTTACACCTATTATTATTGAGAGATGCCTCTCTTTTTGGTAGTGAA 615

Qy        225 TrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArg 244
Db        616 TTTGGGCTTACATCCCAAGAAATTCAACGTTATTATGAGCGCCAAGTGGAAAAAACGAGA 675

Qy        245 GluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSer 264
Db        676 GAATATTCTGATTATTGCGCAAGATGGTATAATACGGGTTTAAATAATTTGAGAGGGACA 735

Qy        265 AsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeu 284
Db        736 AATGCTGAAAGTTGGTTGCGATATAATCAATTCCGTAGAGACTTAACGCTAGGAGTATTA 795

Qy        285 AspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLysThr 304
Db        796 GATCTAGTGGCACTATTCCCAAGCTATGACACGCGTGTTTATCCAATGAATACCAGTGCT 855

Qy        305 GluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGly 324
Db        856 CAATTAACAAGAGAAATTTATACAGATCCAATTGGGAGAACAAATGCA---CCTTCAGGA 912

Qy        325 GlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsn 344
Db        913 TTTGCAAGTACGAATTGGTTTAATAATAATGCACCATCGTTTTCTGCCATAGAGGCTGCC 972

Qy        345 AlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThr 364
Db        973 GTTATTAGGCCTCCGCATCTACTTGATTTTCCAGAACAGCTTACAATTTTCAGC---GTA 1029

Qy        365 LeuGlyAsnMetSerAspValArg------AsnIleTrpGlyGlyHisThrLeuValGlu 382
Db       1030 TTAAGTCGATGGAGTAATACTCAATATATGAATTACTGGGTGGGACATAGACTTGAATCG 1089

Qy        383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp---SerIleThr 401
Db       1090 CGAACAATAAGGGGGTCATTAAGTACCTGGACACACGGAAATACCAATACTTCTATTAAT 1149

Qy        402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
Db       1150 CCTGTAACA---TTACAGTTCACATCTCGAGACGTTTATAGAACAGAATCATTTGCAGGG 1206

Qy        422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
Db       1207 ATAAATATACTTCTAACTACTCCTGTGAAT------------------GGAGTACCTTGG 1248

Qy        442 ValIlePheAsnThrSerAsn---IleAsnAsnValProGlySerLeuArgTyrGluVal 460
Db       1249 GCTAGATTTAATTGGAGAAATCCCCTGAATTCTCTTAGAGGTAGCCTTCTCTATACTATA 1308

Qy        461 ProAlaAsn------------------LeuProSerGlnThrIleLeuSerGluLeuPro 474
Db       1309 GGGTATACTGGAGTGGGGACACAACTATTTGATTCAGAAACT---------GAATTACCA 1359

Qy        475 GlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSer 494
Db       1360 CCAGAAACAACAGAACGACCAAATTATGAATCTTACAGTCATAGATTATCTAATATA--- 1416

Qy        495 AsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrp 514
Db       1417 ---------------AGACTAATATCAGGAAACACTTTGAGAGCACCAGTATATTCTTGG 1461

Qy        515 AlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAsp 534
Db       1462 ACGCACCGTAGTGCAGATCGTACAAATACCATTAGTTCAGATAGCATAACACAAATACCA 1521

Qy        535 AlaValLysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGlyProThr 551
Db       1522 TTGGTAAAATCATTCAACCTTAATTCAGGTACCTCTGTAGTCAGTGGCCCAGGATTTACA 1581

Qy        552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
Db       1582 GGAGGGGATATAATCCGAACTAACGTTAATGGTAGTGTACTAAGTATGGGTCTTAATTTT 1641

Qy        572 ArgGlnThrSer---------TyrArgIleArgLeuArgTyrAlaCysLeuValThrHis 588
Db       1642 AATAATACATCATTACAGCGGTATCGCGTGAGAGTTCGTTATGCTGCTTCTCAAACAATG 1701

Qy        589 GlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCys 608
Db       1702 ------GTCCTGAGGGTAACTGTCGGAGGGAGTACTACTTTTGATCAAGGATTCCCTAGT 1755

Qy        609 SerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAsp 628
Db       1756 ACTATGAGTGCA---------AATGAGTCTTTGACATCTCAATCATTTAGATTTGCAGAA 1806

Qy        629 ValPro------------------GlyIlePheThrProSerIleAsnProLeuIleArg 642
Db       1807 TTTCCTGTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATAAGTAATAAT 1866

Qy        643 TyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db       1867 GCAGGTAGACAAACGTTT---------CACTTTGATAAAATTGAATTCATTCCAATTACT 1917

Qy        663 ---ThrPheProAsnGln---SerLeuGluLysArgGluGlnGluValAsnAspLeuPhe 680
Db       1918 GCAACCTTCGAAGCAGAATATGATTTAGAAAGAGCGCAAGAGGCGGTGAATGCTCTGTTT 1977

Qy        681 IleAsn 682
Db       1978 ACTAAT 1983
```

```
RESULT 12
US-09-053-549-1
; Sequence 1, Application US/09053549
; Patent No. 6121521
;  GENERAL INFORMATION:
;    APPLICANT:  Desai, Nalini
;    TITLE OF INVENTION:  No. 6121521el Insecticidal Protein and Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6121521artis Corporation
;      STREET:  3054 Cornwallis Rd.
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;     APPLICATION NUMBER:  US/09/053,549
;     FILING DATE:  01-APR-1998
;     CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;     NAME:  Pace, Gary M.
;     REGISTRATION NUMBER:  40,403
;     REFERENCE/DOCKET NUMBER:  CGC 1995
;    TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  919-541-8582
;     TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH:  8854 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  11..3694
;     OTHER INFORMATION:  /product= "hyFLIB protein"
;    FEATURE:
;     NAME/KEY:  misc_feature
;     LOCATION:  8854
;     OTHER INFORMATION:  /note= "Sequence of pCIB5520
; Patent No. 6121521
;     OTHER INFORMATION:  containing coding sequence for hyFLIB protein"
US-09-053-549-1

Alignment Scores:
Pred. No.:              3.99e-84      Length:         8854
Score:                  892.00        Matches:        250
Percent Similarity:     50.2%         Conservative:   94
Best Local Similarity:  36.5%         Mismatches:     281
Query Match:            24.8%         Indels:         60
DB:                     3             Gaps:           24

US-10-782-096-2 (1-682) x US-09-053-549-1 (1-8854)

Qy      34 AspProGlnMetThrMetArgAsnThrAsnTyrLysGluTrpLeuAsnMetCysAspSer 53
Db       2 GATCCAACAATGACCAGCAACCGCAAGAACGAGAACGAGATCATCAACGCCGTG---AGC 58

Qy      54 AsnThrGlnPheIleGlyAspIleSerThrTyrSerSerProGluAlaAlaLeuSerVal 73
Db      59 AACCACAGCGCCCAGATGGACCTGCTGCCCGACGCCCGCATCGAGGACAGCCTGTGCATC 118

Qy      74 ArgAsp------------------------------------AlaValLeuThrGlyIleAsn 82
Db     119 GCCGAGGGCAACAACATCGACCCCTTCGTGAGCGCCAGCACCGTGCAGACCGGCATCAAC 178

Qy      83 SerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIle 102
Db     179 ATCGCCGGCCGCATCCTGGGCGTGCTGGGCGTGCCCTTCGCCGGCCAGCTGGCTAGCTTC 238

Qy     103 IleSerArgLeuIleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMet 121
Db     239 TACAGCTTCCTGGTGGGCGAGCTGTGGCCCCGCGGCCGCGACCAGTGGGAAATCTTCCTG 298

Qy     122 ValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeu 141
Db     299 GAGCACGTGGAGCAGCTGATCAACCAGCAGATCACCGAGAACGCCCGCAACACCGCTCTT 358

Qy     142 ArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrp 161
Db     359 GCCCGCCTGCAGGGTCTGGGCGACAGCTTCCGCGCCTACCAGCAGAGCCTGGAGGACTGG 418

Qy     162 LeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAlaIleVal 180
Db     419 CTGGAGAACCGCGACGACGCCCGCACCCGCAGCGTGCTGTACACCCAGTACATCGCCCTG 478

Qy     181 AspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeu 200
Db     479 GAGCTGGACTTCCTGAACGCCATGCCCCTGTTCGCCATTCGAAACCAGGAGGTGCCCCTG 538

Qy     201 LeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyr 220
Db     539 CTGATGGTGTACGCCCAGGCCGCCAACCTGCACCTGCTGCTGCTGCGCGACGCCAGCCTG 598

Qy     221 PheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGln 240
Db     599 TTCGGCAGCGAGTTCGGCCTGACCAGCCAGGAGATCCAGCGCTACTACGAGCGCCAGGTG 658

Qy     241 GlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGln 260
Db     659 GAGCGCACCCGCGACTACAGCGACTACTGCGTGGAGTGGTACAACACCGGCCTGAACAGC 718

Qy     261 PheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThr 280
Db     719 TTAAGGGGCACCAACGCCGCCAGCTGGGTGCGCTACAACCAGTTCCGCCGCGACCTGACC 778

Qy     281 LeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeu 300
Db     779 CTGGGCGTGCTGGACCTGGTGGCCCTGTTCCCCAGCTACGACACCCGCACCTACCCCATC 838

Qy     301 AlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyVal 320
Db     839 AACACCAGCGCCCAGCTGACCCGCGAGGTGTACACCGACGCCATCGGCGCCACCGGCGTG 898

Qy     321 LeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAla 340
Db     899 ---------AACATGGCCAGCATGAACTGGTACAACAACAACGCCCCCAGCTTCAGCGCC 949

Qy     341 MetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheVal 360
Db     950 ATCGAGGCCGCCGCCATCCGCAGCCCCCACCTGCTGGACTTCCTGGAGCAGCTGACCATC 1009

Qy     361 TyrThrArgThrLeuGlyAsnMetSerAspValArgAsnIle------TrpGlyGlyHis 378
Db    1010 TTCAGCGCCAGC---AGCCGCTGGAGCAACACCCGCCACATGACCTACTGGCGCGGCCAC 1066

Qy     379 ThrLeuValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp 398
Db    1067 ACCATCCAGTCTAGACCCATCGGCGGCGGCCTGAACACCAGCACCCACGGCGCCACCAAC 1126

Qy     399 ---SerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGlu 417
Db    1127 ACCACCATCAACCCCGTGACC---CTGCGCTTCGCCTCCCGAGACGTCTACCGCACCGAG 1183

Qy     418 SerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyr 437
Db    1184 AGCTACGCCGGCGTGCTGCTGTGGGGCATC------------TACCTGGAGCCCATCCAC 1231

Qy     438 GlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValPro------GlySer 455
Db    1232 GGCGTGCCCACCGTGCGCTTCAACTTCACCAACCCCCAGAACATCAGCGACCGCGGCACC 1291

Qy     456 LeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu---------SerGlu 472
Db    1292 GCCAACTACAGCCAGCCCTACGAGAGCCCCGGGTTGCAGCTGAAGGACAGCGAGACCGAG 1351

Qy     473 LeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyr 492
Db    1352 CTGCCCCCCGAGACCACCGAGCGCCCCAACTACGAGAGCTACAGCCACCGCCTGAGCCAC 1411

Qy     493 IleSerAsnPheAspAlaArgArgSerSerGlyGlyIleValSerLeuLeuThrPhe 512
Db    1412 ATCGGCATCATCTTGCAGAGCCGC------------------GTGAACGTGCCCGTGTAC 1453

Qy     513 GlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGln 532
Db    1454 AGCTGGACCCACCGCAGCGCCGACCGCACCAACACCATCGGCCCCAACCGCATCACCCAG 1513

Qy     533 IleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGly 549
```

```
Db       1514 ATCCCCATGGTGAAGGCCAGCGAGCTGCCCCAGGGCACCACCGTGGTGCGCGGCCCCGGC 1573

Qy        550 ProThrGlyGlyAsnLeuValLysValSerAspSerTrpHis------------SerLeu 565
Db       1574 TTCACCGGCGGCGACATCCTGCGCCGCACCAACACCGGCGGCTTCGGCCCCATCCGCGTG 1633

Qy        566 LysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeu 585
Db       1634 ACCGTGAACGGCCCCCTGACCCAG---CGCTACCGCATCGGCTTCCGCTACGCCAGCACC 1690

Qy        586 ValThrHisGlyAspAlaIlePheValGluHisSerGlySerSer---HisIleValSer 604
Db       1691 GTGGACTTCGAC------TTCTTCGTGAGCCGCGGCGGCACCACCGTGAACAACTTCCGC 1744

Qy        605 PhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPhe 624
Db       1745 TTCCTGCGCACCATGAACAGC------------GGCGACGAGCTGAAGTACGGCAACTTC 1792

Qy        625 ------ArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeuIleArg 642
Db       1793 GTGCGCCGCGCCTTCACCACCCCCTTCACCTTCACC---CAGATCCAGAACATCATCCGC 1849

Qy        643 TyrArgThrGlnSerPhe---------GlyThrHisAlaIleAspLysPheGluPheIle 659
Db       1850 ACCAGCATCCAGGGCCTGAGCGGCAACGGCGAGGTGTACATCGACAAGATCGAGATCATC 1909

Qy        660 ProLeuAsn---ThrPheProAsnGln---SerLeuGluLysArgGluGlnGluValAsn 677
Db       1910 CCCGTGACTGCCACCTTCGAGGCCGAGTATGACCTGGAGCGCGCCCAGGAGGCCGTGAAC 1969

Qy        678 AspLeuPheIleAsn 682
Db       1970 GCCCTGTTCACCAAC 1984
```

RESULT 13
US-09-178-252-22
; Sequence 22, Application US/09178252
; Patent No. 6218188
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2
; CURRENT APPLICATION NUMBER: US/09/178,252
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 60/065,215
; EARLIER FILING DATE: 1997-11-12
; EARLIER APPLICATION NUMBER: 60/076,445
; EARLIER FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 3558
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-178-252-22

Alignment Scores:

Pred. No.:	1.95e-84	Length:	3558
Score:	889.00	Matches:	251
Percent Similarity:	50.7%	Conservative:	104
Best Local Similarity:	35.9%	Mismatches:	273
Query Match:	24.7%	Indels:	72
DB:	3	Gaps:	24

US-10-782-096-2 (1-682) x US-09-178-252-22 (1-3558)

```
Qy          7 LysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyr 26
Db         16 AAGAACGAGAACGAGATCATCAACGCTCTTTCTATCCCAGCT---GTTTCTAACCATTCT 72

Qy         27 ProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGlu 46
Db         73 GCTCAGATGAACCTTTCTACTGATGCTAGAATC--------------------GAGGAT 111

Qy         47 TrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSer 66
Db        112 TCTCTTTGCATTGCTGAGGGAAAC---------------AACATTGATCCATTCGTTTCT 156

Qy         67 ProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThr 86
Db        157 --------------------GCTTCTACTGTTCAAACTGGAATCAACATTGCTGGAAGA 195

Qy         87 IleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeu 106
Db        196 ATCCTTGGAGTTCTTGGAGTTCCATTCGCTGGACAGATTGCTTCTTTCTACTCTTTCCTT 255

Qy        107 IleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMetValLeuValGlu 125
Db        256 GTTGGAGAGCTTTGGCCTAGGGGAAGAGATCCTTGGGAGATCTTCCTTGAGCATGTTGAG 315

Qy        126 GluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGlu 145
Db        316 CAGTTGATTCGTCAACAAGTTACTGAGAACACTAGAGATACTGCTCTTGCTAGACTTCAA 375

Qy        146 GlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLys 165
Db        376 GGACTTGGAAACTCTTTCAGAGCTTACCAACAATCTCTTGAGGATTGGCTTGAGAACAGA 435

Qy        166 AsnAspAspAsnArgArgAla---LeuValThrGlnTyrAlaIleValAspAsnPhePhe 184
Db        436 GATGATGCTAGAACTAGATCTGTGTTGTACACTCAGTACATTGCTCTTGAGCTTGACTTC 495

Qy        185 GluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyr 204
Db        496 TTGAACGCTATGCCATTGTTCGCTATCAGAAACCAAGAGGTTCCACTTCTCATGGTGTAC 555

Qy        205 AlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGln 224
Db        556 GCTCAAGCTGCTAACCTTCATCTTCTTCTTCTTAGAGATGCTAGCTTGTTCGGATCTGAG 615

Qy        225 TrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArg 244
Db        616 TTCGGACTTACTTCTCAAGAGATTCAAAGATACTACGAGAGACAAGTTGAGAAGACTAGA 675

Qy        245 GluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSer 264
Db        676 GAGTACTCTGACTACTGCGCTAGGTGGTACAACACTGGATTGAACAACCTTAGAGGAACT 735

Qy        265 AsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeu 284
Db        736 AACGCTGAGTCTTGGCTTAGATACAACCAGTTCAGAAGAGATCTTACTCTTGGAGTTCTT 795

Qy        285 AspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLysThr 304
Db        796 GATCTTGTTGCCTTGTTCCCATCTTACGATACTAGAGTGTACCCTATGAACACTTCTGCT 855

Qy        305 GluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGly 324
Db        856 CAACTTACTAGAGAGATCTACACTGATCCAATCGGAAGAACTAACGCT---CCATCTGGA 912

Qy        325 GlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsn 344
Db        913 TTCGCTTCTACTAACTGGTTCAACAACAACGCTCCATCTTTCTCTGCTATCGAGGCTGCA 972

Qy        345 AlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThr 364
Db        973 GTGATCAGACCACCACATCTTCTTGACTTCCCAGAGCAACTTACTATCTTCTCT---GTT 1029

Qy        365 LeuGlyAsnMetSerAspValArg------AsnIleTrpGlyGlyHisThrLeuValGlu 382
Db       1030 CTTTCTAGATGGTCTAACACTCAGTACATGAACTACTGGGTTGGACATAGACTTGAGTCT 1089
```

```
Qy        383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp---SerIleThr 401
Db       1090 AGAACTATCAGAGGATCTCTTTCTACTTCTACTCATGGAAACACTAACACTTCTATCAAC 1149

Qy        402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
Db       1150 CCAGTTACT---CTTCAGTTCACTTCTAGAGATGTGTACAGAACTGAGTCTTTCGCTGGA 1206

Qy        422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
Db       1207 ATCAACATTCTTCTTACTACTCCAGTGAAC------------------GGAGTTCCTTGG 1248

Qy        442 ValIlePheAsnThrSerAsn---IleAsnAsnValProGlySerLeuArgTyrGluVal 460
Db       1249 GCTAGATTCAACTGGAGAAACCCATTGAACTCTCTTAGAGGTTCCTTGTTGTACACCATT 1308

Qy        461 ProAlaAsn------------------LeuProSerGlnThrIleLeuSerGluLeuPro 474
Db       1309 GGATACACTGGAGTTGGTACCCAGTTGTTCGATTCTGAGACT---------GAGCTTCCA 1359

Qy        475 GlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSer 494
Db       1360 CCAGAGACTACTGAGAGACCAAACTACGAGTCTTACTCTCATAGACTTTCTAACATT--- 1416

Qy        495 AsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrp 514
Db       1417 ---------------CGTTTGATCTCTGGAAACACTCTTAGAGCTCCAGTGTACTCTTGG 1461

Qy        515 AlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAsp 534
Db       1462 ACTCATAGATCTGCTGATAGAACTAACACCATCTCTTCTGATTCTATCACTCAGATTCCA 1521

Qy        535 AlaValLysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGlyProThr 551
Db       1522 CTTGTGAAGTCTTTCAACTTGAACTCTGGAACTTCTGTTGTTTCTGGACCAGGATTCACT 1581

Qy        552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
Db       1582 GGAGGAGACATCATCAGAACTAACGTGAACGGATCTGTTCTTTCTATGGGATTGAACTTC 1641

Qy        572 ArgGlnThrSer---------TyrArgIleArgLeuArgTyrAlaCysLeuValThrHis 588
Db       1642 AACAACACTTCTCTTCAAAGATACAGAGTTAGAGTTAGATACGCTGCTTCTCAAACTATG 1701

Qy        589 GlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCys 608
Db       1702 ------GTTCTTAGAGTTACTGTTGGAGGATCTACTACTTTCGATCAAGGATTCCCATCT 1755

Qy        609 SerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAsp 628
Db       1756 ACTATGTCTGCT---------AACGAGTCTCTTACTTCTCAATCTTTCAGATTCGCTGAG 1806

Qy        629 ValPro------------------GlyIlePheThrProSerIleAsnProLeuIleArg 642
Db       1807 TTCCCAGTTGGAATCTCTGCTTCTGGATCTCAAACTGCTGGAATCTCTATCTCTAACAAC 1866

Qy        643 TyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db       1867 GCTGGAAGACAAACTTTC---------CACTTCGACAAGATTGAGTTCATTCCAATCACT 1917

Qy        663 ---ThrPheProAsnGlnSer---LeuGluLysArgGluGlnGluValAsnAspLeuPhe 680
Db       1918 GCTACTCTCGAGGCAGAGTCTGACTTGGAAAGAGCACAGAAGGCGGTGAATGCTCTGTTC 1977
```

RESULT 14
US-09-826-660-22
; Sequence 22, Application US/09826660
; Patent No. 6673990
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 3558
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-826-660-22

Alignment Scores:

Pred. No.:	1.95e-84	Length:	3558
Score:	889.00	Matches:	251
Percent Similarity:	50.7%	Conservative:	104
Best Local Similarity:	35.9%	Mismatches:	273
Query Match:	24.7%	Indels:	72
DB:	3	Gaps:	24

US-10-782-096-2 (1-682) x US-09-826-660-22 (1-3558)

```
Qy          7 LysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyr 26
Db         16 AAGAACGAGAACGAGATCATCAACGCTCTTTCTATCCCAGCT---GTTTCTAACCATTCT 72

Qy         27 ProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGlu 46
Db         73 GCTCAGATGAACCTTTCTACTGATGCTAGAATC---------------------GAGGAT 111

Qy         47 TrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSer 66
Db        112 TCTCTTTGCATTGCTGAGGGAAAC---------------AACATTGATCCATTCGTTTCT 156

Qy         67 ProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThr 86
Db        157 ---------------------GCTTCTACTGTTCAAACTGGAATCAACATTGCTGGAAGA 195

Qy         87 IleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeu 106
Db        196 ATCCTTGGAGTTCTTGGAGTTCCATTCGCTGGACAGATTGCTTCTTTCTACTCTTTCCTT 255

Qy        107 IleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMetValLeuValGlu 125
Db        256 GTTGGAGAGCTTTGGCCTAGGGGAAGAGATCCTTGGGAGATCTTCCTTGAGCATGTTGAG 315

Qy        126 GluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGlu 145
Db        316 CAGTTGATTCGTCAACAAGTTACTGAGAACACTAGAGATACTGCTCTTGCTAGACTTCAA 375

Qy        146 GlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLys 165
Db        376 GGACTTGGAAACTCTTTCAGAGCTTACCAACAATCTCTTGAGGATTGGCTTGAGAACAGA 435

Qy        166 AsnAspAspAsnArgArgAla---LeuValThrGlnTyrAlaIleValAspAsnPhePhe 184
Db        436 GATGATGCTAGAACTAGATCTGTGTTGTACACTCAGTACATTGCTCTTGAGCTTGACTTC 495

Qy        185 GluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyr 204
Db        496 TTGAACGCTATGCCATTGTTCGCTATCAGAAACCAAGAGGTTCCACTTCTCATGGTGTAC 555

Qy        205 AlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGln 224
Db        556 GCTCAAGCTGCTAACCTTCATCTTCTTCTTCTTAGAGATGCTAGCTTGTTCGGATCTGAG 615
```

```
Qy       225 TrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArg 244
Db       616 TTCGGACTTACTTCTCAAGAGATTCAAAGATACTACGAGAGACAAGTTGAGAAGACTAGA 675

Qy       245 GluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSer 264
Db       676 GAGTACTCTGACTACTGCGCTAGGTGGTACAACACTGGATTGAACAACCTTAGAGGAACT 735

Qy       265 AsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeu 284
Db       736 AACGCTGAGTCTTGGCTTAGATACAACCAGTTCAGAAGAGATCTTACTCTTGGAGTTCTT 795

Qy       285 AspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLysThr 304
Db       796 GATCTTGTTGCCTTGTTCCCATCTTACGATACTAGAGTGTACCCTATGAACACTTCTGCT 855

Qy       305 GluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGly 324
Db       856 CAACTTACTAGAGAGATCTACACTGATCCAATCGGAAGAACTAACGCT---CCATCTGGA 912

Qy       325 GlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsn 344
Db       913 TTCGCTTCTACTAACTGGTTCAACAACAACGCTCCATCTTTCTCTGCTATCGAGGCTGCA 972

Qy       345 AlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThr 364
Db       973 GTGATCAGACCACCACATCTTCTTGACTTCCCAGAGCAACTTACTATCTTCTCT---GTT 1029

Qy       365 LeuGlyAsnMetSerAspValArg------AsnIleTrpGlyGlyHisThrLeuValGlu 382
Db      1030 CTTTCTAGATGGTCTAACACTCAGTACATGAACTACTGGGTTGGACATAGACTTGAGTCT 1089

Qy       383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp---SerIleThr 401
Db      1090 AGAACTATCAGAGGATCTCTTTCTACTTCTACTCATGGAAACACTAACACTTCTATCAAC 1149

Qy       402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
Db      1150 CCAGTTACT---CTTCAGTTCACTTCTAGAGATGTGTACAGAACTGAGTCTTTCGCTGGA 1206

Qy       422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
Db      1207 ATCAACATTCTTCTTACTACTCCAGTGAAC------------------GGAGTTCCTTGG 1248

Qy       442 ValIlePheAsnThrSerAsn---IleAsnAsnValProGlySerLeuArgTyrGluVal 460
Db      1249 GCTAGATTCAACTGGAGAAACCCATTGAACTCTCTTAGAGGTTCCTTGTTGTACACCATT 1308

Qy       461 ProAlaAsn------------------LeuProSerGlnThrIleLeuSerGluLeuPro 474
Db      1309 GGATACACTGGAGTTGGTACCCAGTTGTTCGATTCTGAGACT---------GAGCTTCCA 1359

Qy       475 GlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSer 494
Db      1360 CCAGAGACTACTGAGAGACCAAACTACGAGTCTTACTCTCATAGACTTTCTAACATT--- 1416

Qy       495 AsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrp 514
Db      1417 ---------------CGTTTGATCTCTGGAAACACTCTTAGAGCTCCAGTGTACTCTTGG 1461

Qy       515 AlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAsp 534
Db      1462 ACTCATAGATCTGCTGATAGAACTAACACCATCTCTTCTGATTCTATCACTCAGATTCCA 1521

Qy       535 AlaValLysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGlyProThr 551
Db      1522 CTTGTGAAGTCTTTCAACTTGAACTCTGGAACTTCTGTTGTTTCTGGACCAGGATTCACT 1581

Qy       552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
Db      1582 GGAGGAGACATCATCAGAACTAACGTGAACGGATCTGTTCTTTCTATGGGATTGAACTTC 1641

Qy       572 ArgGlnThrSer---------TyrArgIleArgLeuArgTyrAlaCysLeuValThrHis 588
Db      1642 AACAACACTTCTCTTCAAAGATACAGAGTTAGAGTTAGATACGCTGCTTCTCAAACTATG 1701

Qy       589 GlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCys 608
Db      1702 ------GTTCTTAGAGTTACTGTTGGAGGATCTACTACTTTCGATCAAGGATTCCCATCT 1755

Qy       609 SerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAsp 628
Db      1756 ACTATGTCTGCT---------AACGAGTCTCTTACTTCTCAATCTTTCAGATTCGCTGAG 1806

Qy       629 ValPro------------------GlyIlePheThrProSerIleAsnProLeuIleArg 642
Db      1807 TTCCCAGTTGGAATCTCTGCTTCTGGATCTCAAACTGCTGGAATCTCTATCTCTAACAAC 1866

Qy       643 TyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db      1867 GCTGGAAGACAAACTTTC---------CACTTCGACAAGATTGAGTTCATTCCAATCACT 1917

Qy       663 ---ThrPheProAsnGlnSer---LeuGluLysArgGluGlnGluValAsnAspLeuPhe 680
Db      1918 GCTACTCTCGAGGCAGAGTCTGACTTGGAAAGAGCACAGAAGGCGGTGAATGCTCTGTTC 1977
```

RESULT 15
US-08-377-690-1
; Sequence 1, Application US/08377690
; Patent No. 5628995
; GENERAL INFORMATION:
; APPLICANT: PEFEROEN, Marnix
; APPLICANT: JANSENS, Stefan
; APPLICANT: DENOLF, Peter
; TITLE OF INVENTION: CONTROL OF OSTRINIA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: The George Mason Bldg., Washington & Prince
; STREET: Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,690
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/164,781
; FILING DATE:
; APPLICATION NUMBER: US 07/938,362
; FILING DATE: 31-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-039
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4074 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: entomocidus HD 110
; FEATURE:

```
;       NAME/KEY:  CDS
;       LOCATION:  186..3872
;       OTHER INFORMATION:  /note= "PROPERTIES:  CryIB is toxic to
;       OTHER INFORMATION:  Ostrinea nubilalis (among others)"
US-08-377-690-1

Alignment Scores:
Pred. No.:                4.48e-84      Length:        4074
Score:                    886.50        Matches:       249
Percent Similarity:       50.1%         Conservative:  102
Best Local Similarity:    35.5%         Mismatches:    277
Query Match:              24.6%         Indels:        73
DB:                       2             Gaps:          25

US-10-782-096-2 (1-682) x US-08-377-690-1 (1-4074)

Qy        7 LysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyr 26
Db      201 AAAAATGAGAATGAAATTATAAATGCTGTATCGAATCATTCCGCACAAATGGATCTATTA 260

Qy       27 ProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGlu 46
Db      261 CCA-------------------------------------GATGCTCGTATTGAGGAT 281

Qy       47 TrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSer 66
Db      282 AGCTTGTGTATAGCCGAGGGGAAC---------------AATATTGATCCATTTGTTAGC 326

Qy       67 ProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThr 86
Db      327 ---------------------GCATCAACAGTCCAAACGGGTATTAACATAGCTGGTAGA 365

Qy       87 IleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeu 106
Db      366 ATACTAGGCGTATTGGGCGTACCGTTTGCTGGACAACTAGCTAGTTTTTATAGTTTTCTT 425

Qy      107 IleGlyIleLeuTrp---AlaGlyProAspProPheGluAlaLeuMetValLeuValGlu 125
Db      426 GTTGGTGAATTATGGCCCCGCGGCAGAGATCAGTGGGAAATTTTCCTAGAACATGTCGAA 485

Qy      126 GluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGlu 145
Db      486 CAACTTATAAATCAACAAATAACAGAAAATGCTAGGAATACGGCTCTTGCTCGATTACAA 545

Qy      146 GlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLys 165
Db      546 GGTTTAGGAGATTCCTTCAGAGCCTATCAACAGTCACTTGAAGATTGGCTAGAAAACCGT 605

Qy      166 AsnAspAspAsnArgArgAlaLeuVal---ThrGlnTyrAlaIleValAspAsnPhePhe 184
Db      606 GATGATGCAAGAACGAGAAGTGTTCTTCATACCCAATATATAGCTTTAGAACTTGATTTT 665

Qy      185 GluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeuProValTyr 204
Db      666 CTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAAGTTCCATTATTGATGGTATAT 725

Qy      205 AlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGln 224
Db      726 GCTCAAGCTGCAAATTTACACCTATTATTATTGAGAGATGCCTCTCTTTTTGGTAGTGAA 785

Qy      225 TrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArg 244
Db      786 TTTGGGCTTACATCGCAGGAAATTCAACGCTATTATGAGCGCCAAGTGGAACGAACGAGA 845

Qy      245 GluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSer 264
Db      846 GATTATTCCGACTATTGCGTAGAATGGTATAATACAGGTCTAAATAGCTTGAGAGGGACA 905

Qy      265 AsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeuThrValLeu 284
Db      906 AATGCCGCAAGTTGGGTACGGTATAATCAATTCCGTAGAGATCTAACGTTAGGAGTATTA 965

Qy      285 AspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAlaValLysThr 304
Db      966 GATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCAATAAATACGAGTGCT 1025

Qy      305 GluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeuGluSerGly 324
Db     1026 CAGTTAACAAGAGAAGTTTATACAGACGCAATTGGAGCAACAGGGGTA---------AAT 1076

Qy      325 GlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMetGluAsnAsn 344
Db     1077 ATGGCAAGTATGAATTGGTATAATAATAATGCACCTTCGTTCTCTGCCATAGAGGCTGCG 1136

Qy      345 AlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyrThrArgThr 364
Db     1137 GCTATCCGAAGCCCGCATCTACTTGATTTTCTAGAACAACTTACAATTTTTAGCGCTTCA 1196

Qy      365 LeuGlyAsnMetSerAspValArgAsnIle------TrpGlyGlyHisThrLeuValGlu 382
Db     1197 ---TCACGATGGAGTAATACTAGGCATATGACTTATTGGCGGGGCACACGATTCAATCT 1253

Qy      383 AsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAsp---SerIleThr 401
Db     1254 CGGCCAATAGGAGGCGGATTAAATACCTCAACGCATGGGGCTACCAATACTTCTATTAAT 1313

Qy      402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
Db     1314 CCTGTAACA---TTACGGTTCGCATCTCGAGACGTTTATAGGACTGAATCATATGCAGGA 1370

Qy      422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
Db     1371 GTGCTTCTATGGGGAATT------------TACCTTGAACCTATTCATGGTGTCCCTACT 1418

Qy      442 ValIlePheAsnThrSerAsnIleAsnAsnValPro------GlySerLeuArgTyrGlu 459
Db     1419 GTTAGGTTTAATTTTACGAACCCTCAGAATATTTCTGATAGAGGTACCGCTAACTATAGT 1478

Qy      460 ValProAlaAsnLeuProSerGlnThrIleLeu---------SerGluLeuProGlyLys 476
Db     1479 CAACCTTATGAGTCACCTGGGCTTCAATTAAAAGATTCAGAAACTGAATTACCACCAGAA 1538

Qy      477 AspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPhe 496
Db     1539 ACAACAGAACGACCAAATTATGAATCTTACAGTCACAGGTTATCTCATATAGGTATAATT 1598

Qy      497 AspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHis 516
Db     1599 TTACAATCCAGG------------------GTGAATGTACCGGTATATTCTTGGACGCAT 1640

Qy      517 ThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaVal 536
Db     1641 CGTAGTGCAGATCGTACGAATACGATTGGACCAAATAGAATCACCCAAATCCCAATGGTA 1700

Qy      537 LysGlyTrpGlyGlyAsnIleGlyPheValIle---------ProGlyProThrGlyGly 553
Db     1701 AAAGCATCCGAACTTCCTCAAGGTACCACTGTTGTTAGAGGACCAGGATTTACTGGTGGG 1760

Qy      554 AsnLeuValLysValSerAspSerTrpHis------------SerLeuLysValGlnAla 569
Db     1761 GATATTCTTCGAAGAACGAATACTGGTGGATTTGGACCGATAAGAGTAACTGTTAACGGA 1820

Qy      570 ProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGly 589
Db     1821 CCATTAACACAA---AGATATCGTATAGGATTCCGCTATGCTTCAACTGTAGATTTTGAT 1877

Qy      590 AspAlaIlePheValGluHisSerGlySerSer---HisIleValSerPhePheAspCys 608
Db     1878 ------TTCTTTGTATCACGTGGAGGTACTACTGTAAATAATTTTAGATTCCTACGTACA 1931

Qy      609 SerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPhe------ArgTyr 626
Db     1932 ATGAACAGT------------GGAGACGAACTAAAATACGGAAATTTTGTGAGACGTGCT 1979

Qy      627 IleAspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArgThrGln 646
```

```
Db       1980 TTTACTACACCTTTTACTTTTACA---CAAATTCAAGATATAATTCGAACGTCTATTCAA 2036

Qy        647 SerPhe---------GlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                             |||           |||||||||   |||   ||||||:::
Db       2037 GGCCTTAGTGGAAATGGGGAAGTGTATATAGATAAAATTGAAATTATTCCAGTTACTGCA 2096

Qy        663 ThrPheProAsnGln---SerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              ||||||      :::      ||||||:::   ::::::   ||||||   ||||||
Db       2097 ACCTTCGAAGCAGAATATGATTTAGAAAGAGCGCAAGAGGCGGTGAATGCTCTGTTTACT 2156

Qy        682 Asn 682
              |||
Db       2157 AAT 2159
```

Search completed: January 21, 2006, 04:00:20
Job time : 391 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 20, 2006, 23:03:47 ; Search time 1397 Seconds
                                             (without alignments)
                                             4037.019 Million cell updates/sec

Title:          US-10-782-096-2
Perfect score:  3597
Sequence:       1 MNSYKNKNEYEMLDALRINS..........TFPNQSLEKREQEVNDLFIN 682

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters:      19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10782096/runat_20012006_095216_24424/app_query.fasta_1.839
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10782096_@CGN_1_1_1026_@runat_20012006_095216_24424 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA_Main:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                8:  /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
                10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	948	26.4	3621	5	US-10-032-717-1	Sequence 1, Appli
2	948	26.4	3621	6	US-10-414-637-1	Sequence 1, Appli
3	948	26.4	3621	7	US-10-606-320-1	Sequence 1, Appli
4	948	26.4	3621	8	US-10-746-914-1	Sequence 1, Appli
5	948	26.4	3621	10	US-11-021-115-5	Sequence 5, Appli
6	948	26.4	4874	5	US-10-032-717-27	Sequence 27, Appl
7	948	26.4	4874	6	US-10-414-637-27	Sequence 27, Appl
8	948	26.4	4874	7	US-10-606-320-17	Sequence 17, Appl
9	948	26.4	4874	8	US-10-746-914-17	Sequence 17, Appl
10	940.5	26.1	3504	5	US-10-089-678-2	Sequence 2, Appli
11	940.5	26.1	3690	5	US-10-089-678-3	Sequence 3, Appli
12	926	25.7	2003	5	US-10-032-717-5	Sequence 5, Appli
13	926	25.7	2003	6	US-10-414-637-5	Sequence 5, Appli
14	926	25.7	2010	5	US-10-032-717-9	Sequence 9, Appli
15	926	25.7	2010	5	US-10-032-717-15	Sequence 15, Appl
16	926	25.7	2010	6	US-10-414-637-9	Sequence 9, Appli
17	926	25.7	2010	6	US-10-414-637-15	Sequence 15, Appl
18	926	25.7	2010	7	US-10-606-320-5	Sequence 5, Appli
19	926	25.7	2010	7	US-10-606-320-11	Sequence 11, Appl
20	926	25.7	2010	8	US-10-746-914-5	Sequence 5, Appli
21	926	25.7	2010	8	US-10-746-914-11	Sequence 11, Appl
22	925.5	25.7	2022	7	US-10-606-320-49	Sequence 49, Appl
23	925.5	25.7	2022	7	US-10-606-320-81	Sequence 81, Appl
24	925.5	25.7	2022	8	US-10-746-914-49	Sequence 49, Appl
25	925.5	25.7	2022	8	US-10-746-914-81	Sequence 81, Appl
26	925	25.7	2022	5	US-10-032-717-11	Sequence 11, Appl
27	925	25.7	2022	6	US-10-414-637-11	Sequence 11, Appl
28	925	25.7	2022	7	US-10-606-320-7	Sequence 7, Appli
29	925	25.7	2022	7	US-10-606-320-25	Sequence 25, Appl
30	925	25.7	2022	7	US-10-606-320-29	Sequence 29, Appl
31	925	25.7	2022	7	US-10-606-320-33	Sequence 33, Appl
32	925	25.7	2022	7	US-10-606-320-69	Sequence 69, Appl
33	925	25.7	2022	8	US-10-746-914-7	Sequence 7, Appli
34	925	25.7	2022	8	US-10-746-914-25	Sequence 25, Appl
35	925	25.7	2022	8	US-10-746-914-29	Sequence 29, Appl
36	925	25.7	2022	8	US-10-746-914-33	Sequence 33, Appl
37	925	25.7	2022	8	US-10-746-914-69	Sequence 69, Appl
38	924.5	25.7	2022	7	US-10-606-320-43	Sequence 43, Appl
39	924.5	25.7	2022	7	US-10-606-320-75	Sequence 75, Appl
40	924.5	25.7	2022	8	US-10-746-914-43	Sequence 43, Appl
41	924.5	25.7	2022	8	US-10-746-914-75	Sequence 75, Appl
42	924	25.7	2025	7	US-10-606-320-45	Sequence 45, Appl
43	924	25.7	2025	7	US-10-606-320-77	Sequence 77, Appl
44	924	25.7	2025	8	US-10-746-914-45	Sequence 45, Appl
45	924	25.7	2025	8	US-10-746-914-77	Sequence 77, Appl

ALIGNMENTS

```
RESULT 1
US-10-032-717-1
; Sequence 1, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Nicholas B. Duck
;  APPLICANT:  Xiang Feng
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Lynn E. Sims
;  TITLE OF INVENTION:  Genes Encoding No. US20020151709A1el Proteins With
;  TITLE OF INVENTION:  Pesticidal Activity Against Coleopterans
;  FILE REFERENCE: 35718/237005
;  CURRENT APPLICATION NUMBER: US/10/032,717
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 60/242,838
;  PRIOR FILING DATE: 2000-10-24
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3621)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Cry1218-1
US-10-032-717-1
```

```
Alignment Scores:
Pred. No.:              3.04e-96      Length:        3621
Score:                  948.00        Matches:       248
Percent Similarity:     50.9%         Conservative:  119
Best Local Similarity:  34.4%         Mismatches:    276
Query Match:            26.4%         Indels:        78
DB:                     5             Gaps:          25

US-10-782-096-2 (1-682) x US-10-032-717-1 (1-3621)

Qy      3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db      4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy     23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db     61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy     43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db    121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy     63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db    163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy     82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db    223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy    102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db    283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy    120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db    343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy    140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db    403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTCAA 462

Qy    160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db    463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy    179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db    523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy    199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db    583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy    219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db    643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy    239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db    703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db    763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db    823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db    883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db    943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db    988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy    359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db   1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy    377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db   1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy    396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db   1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy    416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db   1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy    431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db   1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy    450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db   1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy    470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db   1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy    483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db   1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy    503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db   1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy    523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db   1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy    543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db   1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy    559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
Db   1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy    572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db   1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy    592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db   1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy    612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db   1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy    632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644
Db   1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932
```

```
Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::        |||          :::|||:::   |||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||        ::::::   ||||||  ||||||
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
              |||
Db       2053 AAT 2055

RESULT 2
US-10-414-637-1
; Sequence 1, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Nicholas B. Duck
;  APPLICANT:  Xiang Feng
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Lynn E. Sims
;  TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
;  TITLE OF INVENTION:  Pesticidal Activity Against Coleopterans
;  FILE REFERENCE: 35718/237005
;  CURRENT APPLICATION NUMBER: US/10/414,637
;  CURRENT FILING DATE:  2003-04-16
;  PRIOR APPLICATION NUMBER: US/10/032,717
;  PRIOR FILING DATE: 2001-10-23
;  PRIOR APPLICATION NUMBER: 60/242,838
;  PRIOR FILING DATE: 2000-10-24
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3621)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Cry1218-1
US-10-414-637-1

Alignment Scores:
Pred. No.:              3.04e-96      Length:        3621
Score:                  948.00        Matches:       248
Percent Similarity:     50.9%         Conservative:  119
Best Local Similarity:  34.4%         Mismatches:    276
Query Match:            26.4%         Indels:        78
DB:                     6             Gaps:          25

US-10-782-096-2 (1-682) x US-10-414-637-1 (1-3621)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||      |||:::|||||||||||||::::::||||||            ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||         |||||||||   |||      :::   ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                     |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
              |||         |||||||||   :::|||   :::||||||           |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||   :::||||||   ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||              :::         :::|||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||   :::   |||   :::      |||
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||||||||||||         :::||||||   |||   |||:::
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||               |||   :::      ::::::
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::   ||||||   |||:::      |||||||||:::
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                 |||   |||||||||   ||||||||||||||||:::||||||:::||||||
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||               |||   :::   |||   |||
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                       |||      ||||||   |||||||||:::   :::|||      ||||||
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                 :::            ::::::|||:::   ||||||   :::|||:::|||||   :::
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   |||||||||:::         ||||||||||||||||   |||   |||
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||      :::||||||||||||||||||||||||||::::::|||
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||      |||                  ||||||:::         :::|||
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                 :::|||::::::      |||   |||               ::::::      :::
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 |||||||||      |||::::::   :::            :::|||:::   |||
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::                  :::   ||||||   :::      :::|||
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                 ::::::                     |||:::|||   |||         ::::::
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                 |||                  |||      :::|||   |||   |||
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275
```

```
Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 |||||     :::|||:::      |||    |||     :::  :::
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||      :::|||:::|||            |||     |||:::   |||
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                     ||||||||| :::      :::||||||
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                       :::||||||||||||   :::|||::::::      |||      ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                  |||  |||  |||  |||  |||  |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::         ||||||||||||||||  |||||||||  |||      |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||               ||||||  |||||||||:::|||::::::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::   |||    |||:::                             |||   :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::||||||||||||||               ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                 |||   |||             :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::         |||      |||                |||:::   |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                          :::   :::      :::   :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::         |||      :::|||:::   ||||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||      ::::::   ||||||   ||||||
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
              |||
Db       2053 AAT 2055
```

RESULT 3
US-10-606-320-1
; Sequence 1, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/263948
; CURRENT APPLICATION NUMBER: US/10/606,320
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3621)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Cry1218-1
US-10-606-320-1

```
Alignment Scores:
Pred. No.:                3.04e-96      Length:         3621
Score:                    948.00        Matches:        248
Percent Similarity:       50.9%         Conservative:   119
Best Local Similarity:    34.4%         Mismatches:     276
Query Match:              26.4%         Indels:         78
DB:                       7             Gaps:           25

US-10-782-096-2 (1-682) x US-10-606-320-1 (1-3621)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||         |||:::||||||||||||::::::||||||           ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||         |||||||||   |||   :::|||         ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                       |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||      :::|||   :::||||||             |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::      ||||||   :::||||||   ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||                  :::         ::::||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||         ||||||||||||||     :::   |||   :::      |||
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||||||||||             :::||||||   |||   |||:::
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||               |||   :::      ::::::
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::   ||||||   |||:::      |||||||||:::
```

```
Db    523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy    199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218

Db    583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy    219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238

Db    643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy    239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258

Db    703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278

Db    763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298

Db    823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318

Db    883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338

Db    943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358

Db    988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy    359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376

Db   1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy    377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395

Db   1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy    396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415

Db   1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy    416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430

Db   1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy    431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449

Db   1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy    450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469

Db   1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy    470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482

Db   1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy    483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502

Db   1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy    503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522

Db   1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy    523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542

Db   1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy    543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558

Db   1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy    559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571

Db   1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy    572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591

Db   1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy    592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611

Db   1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy    612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631

Db   1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy    632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644

Db   1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy    645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662

Db   1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy    663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681

Db   1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy    682 Asn 682

Db   2053 AAT 2055
```

RESULT 4
US-10-746-914-1
; Sequence 1, Application US/10746914
; Publication No. US20040210963A1
; GENERAL INFORMATION:
; APPLICANT: Albert L. Lu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/268350
; CURRENT APPLICATION NUMBER: US/10/746,914
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3621)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Cry1218-1
US-10-746-914-1

Alignment Scores:

Pred. No.:	3.04e-96	Length:	3621
Score:	948.00	Matches:	248
Percent Similarity:	50.9%	Conservative:	119

```
Best Local Similarity:  34.4%         Mismatches:   276
Query Match:            26.4%         Indels:        78
DB:                     8             Gaps:          25

US-10-782-096-2 (1-682) x US-10-746-914-1 (1-3621)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992
```

```
Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
Db       2053 AAT 2055
```

RESULT 5
US-11-021-115-5
; Sequence 5, Application US/11021115
; Publication No. US20050166284A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Andre
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Herrmann, Rafael
; APPLICANT: Lu, Albert L.
; APPLICANT: McCutchen, Bill F.
; APPLICANT: Presnail, James K.
; APPLICANT: Rice, Janet A.
; APPLICANT: Wong, James F.
; APPLICANT: Yu, Cao-Guo
; TITLE OF INVENTION: Plant Activation of Insect Toxin
; FILE REFERENCE: 035718/285836
; CURRENT APPLICATION NUMBER: US/11/021,115
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: 60/532,185
; PRIOR FILING DATE: 2003-12-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-11-021-115-5

Alignment Scores:

Pred. No.:	3.04e-96	Length:	3621
Score:	948.00	Matches:	248
Percent Similarity:	50.9%	Conservative:	119
Best Local Similarity:	34.4%	Mismatches:	276
Query Match:	26.4%	Indels:	78
DB:	10	Gaps:	25

US-10-782-096-2 (1-682) x US-11-021-115-5 (1-3621)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434
```

```
Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                       :::|||||||||||    :::|||::::::        |||       ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||    :::                   |||   |||   |||   |||   |||   |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||    :::         ||||||||||||||||||   |||||||||   |||         |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::    ||||||                |||||   |||||||||:::|||::::::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::    |||       |||:::                                 |||     :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::             ||||||:::||||||||||||||                  ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                       |||   |||               :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::          |||          |||            |||:::      |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                                :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                   :::          |||          :::|||:::      ||||||||||||:::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::          |||   ||||||          ::::::   ||||||   ||||||
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
              |||
Db       2053 AAT 2055
```

RESULT 6
US-10-032-717-27
; Sequence 27, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 4874
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic DNA 1218-1
US-10-032-717-27

```
Alignment Scores:
Pred. No.:              4.96e-96        Length:         4874
Score:                  948.00          Matches:        248
Percent Similarity:     50.9%           Conservative:   119
Best Local Similarity:  34.4%           Mismatches:     276
Query Match:            26.4%           Indels:         78
DB:                     5               Gaps:           25
```

US-10-782-096-2 (1-682) x US-10-032-717-27 (1-4874)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||         |||:::||||||||||||::::::||||||           ::::::::::::
Db        734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||            |||||||||   |||      :::|||           ::::::|||
Db        791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                              |||:::   |||
Db        851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||         :::|||   :::||||||               |||
Db        893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||   :::||||||   ||||||||||||
Db        953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||            :::            :::|||
Db       1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGAAATT 1072

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||     :::      |||      :::         |||
Db       1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||||||||||            :::||||||   |||   |||:::
Db       1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||                  |||      :::         ::::::
Db       1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::   ||||||   |||:::      |||||||||:::
Db       1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                       |||   |||||||||   |||||||||||||||:::||||||:::||||||
Db       1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                       ||||||   :::|||                  |||   :::   |||   |||
Db       1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                       |||      ||||||   |||||||||:::   :::|||      ||||||
Db       1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                 :::         ::::::|||:::   ||||||   :::|||:::||||||      :::
Db       1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   |||||||||:::         ||||||||||||||||||   |||   |||
```

```
Db       1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||    :::||||||||||||||||||||||||||||:::|||
Db       1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||      |||                  ||||||:::           :::|||
Db       1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                    :::|||::::::      |||   |||         ::::::       :::
Db       1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 ||||||||        |||::::::   :::         :::|||:::   |||
Db       1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::                  :::   ||||||   :::         :::|||
Db       1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                       ::::::                  |||:::|||   |||         ::::::
Db       1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                       |||                  |||      :::|||   |||   |||
Db       1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||         :::|||:::         |||      |||      :::   :::
Db       2006 ---TATATATT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||                  |||      |||:::   |||
Db       2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                     ||||||||   :::         :::||||||
Db       2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                       :::||||||||||||      :::|||:::      |||      ::::::
Db       2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                  |||   |||   |||   |||   |||   |||
Db       2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::      |||||||||||||||   |||||||||   |||      |||
Db       2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||                  ||||||   |||||||||:::|||::::::
Db       2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::      |||   |||:::                           |||   :::
Db       2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::|||||||||||||                  ||||||
Db       2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                    |||   |||         :::   |||
Db       2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::         |||         |||         |||:::      |||
Db       2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                       :::   :::      :::      :::
Db       2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::      |||   ||||||      ::::::   ||||||   ||||||
Db       2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy        682 Asn 682
              |||
Db       2783 AAT 2785
```

RESULT 7
US-10-414-637-27
; Sequence 27, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 4874
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Genomic DNA 1218-1
US-10-414-637-27

Alignment Scores:

Pred. No.:	4.96e-96	Length:	4874
Score:	948.00	Matches:	248
Percent Similarity:	50.9%	Conservative:	119
Best Local Similarity:	34.4%	Mismatches:	276
Query Match:	26.4%	Indels:	78
DB:	6	Gaps:	25

US-10-782-096-2 (1-682) x US-10-414-637-27 (1-4874)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                 |||      |||:::||||||||||||::::::||||||            ::::::::::::
Db        734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
                 ||||||         |||||||||   |||      :::|||         ::::::|||
Db        791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                     |||:::   |||
Db        851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||      :::|||   :::||||||               |||
```

```
Db      893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db     1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db     1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db     1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db     1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db     1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db     1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db     1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db     1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db     1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db     1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db     1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db     1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db     1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db     1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db     1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db     1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy      416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db     1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db     2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db     2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy      470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db     2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db     2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db     2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db     2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db     2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy      559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db     2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db     2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy      592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db     2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy      612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db     2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy      632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db     2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy      645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db     2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy      663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db     2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy      682 Asn 682
Db     2783 AAT 2785
```

RESULT 8
US-10-606-320-17
; Sequence 17, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail

```
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/263948
; CURRENT APPLICATION NUMBER: US/10/606,320
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 4874
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Genomic DNA 1218-1
US-10-606-320-17

Alignment Scores:
Pred. No.:              4.96e-96        Length:         4874
Score:                  948.00          Matches:        248
Percent Similarity:     50.9%           Conservative:   119
Best Local Similarity:  34.4%           Mismatches:     276
Query Match:            26.4%           Indels:         78
DB:                     7               Gaps:           25

US-10-782-096-2 (1-682) x US-10-606-320-17 (1-4874)

Qy      3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db    734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy     23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db    791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy     43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db    851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy     63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db    893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy     82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db    953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy    102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db   1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy    120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db   1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy    140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db   1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy    160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db   1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy    179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db   1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy    199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db   1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy    219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db   1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy    239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db   1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db   1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db   1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db   1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db   1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db   1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy    359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db   1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy    377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db   1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy    396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db   1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy    416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db   1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy    431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db   2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy    450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db   2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy    470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db   2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy    483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db   2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy    503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db   2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy    523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db   2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy    543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
```

```
Db      2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy       559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571

Db      2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy       572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591

Db      2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy       592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611

Db      2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy       612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631

Db      2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy       632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644

Db      2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy       645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662

Db      2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy       663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681

Db      2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy       682 Asn 682

Db      2783 AAT 2785
```

RESULT 9
US-10-746-914-17
; Sequence 17, Application US/10746914
; Publication No. US20040210963A1
; GENERAL INFORMATION:
; APPLICANT: Albert L. Lu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/268350
; CURRENT APPLICATION NUMBER: US/10/746,914
; CURRENT FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 4874
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Genomic DNA 1218-1
US-10-746-914-17

Alignment Scores:

Pred. No.:	4.96e-96	Length:	4874
Score:	948.00	Matches:	248
Percent Similarity:	50.9%	Conservative:	119
Best Local Similarity:	34.4%	Mismatches:	276
Query Match:	26.4%	Indels:	78
DB:	8	Gaps:	25

US-10-782-096-2 (1-682) x US-10-746-914-17 (1-4874)

```
Qy         3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22

Db       734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy        23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42

Db       791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy        43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62

Db       851 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 892

Qy        63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81

Db       893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy        82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101

Db       953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy       102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119

Db      1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy       120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139

Db      1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy       140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159

Db      1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy       160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178

Db      1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy       179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198

Db      1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy       199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218

Db      1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy       219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238

Db      1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy       239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258

Db      1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy       259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278

Db      1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy       279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298

Db      1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy       299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318

Db      1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy       319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338

Db      1673 AACGTGTCTTCAATTGGTTCC-----------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy       339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358

Db      1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777
```

```
Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
              |||||||||        |||:::::   :::      :::||||:::   |||
Db       1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::                 :::  ||||||   :::          :::|||
Db       1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                       :::::::                 |||:::|||   |||      ::::::
Db       1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                 |||                   |||       :::|||   |||   |||
Db       1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||         :::|||:::             |||      |||      :::   :::
Db       2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||                  |||      |||:::   |||
Db       2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                     |||||||||   :::          :::||||||
Db       2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                    :::||||||||||||   :::|||::::::         |||      ::::::
Db       2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                |||   |||   |||   |||   |||   |||
Db       2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::         ||||||||||||||||||   |||      |||      |||
Db       2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||                ||||||   |||||||||:::|||::::::
Db       2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy        559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
              :::   |||      |||:::                              |||      :::
Db       2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::||||||||||||||                  ||||||
Db       2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                 |||      |||      |||                                 |||
Db       2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::      |||      |||         |||:::      |||
Db       2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                          :::   :::      :::      :::
Db       2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::      |||         :::|||:::      ||||||||||||::::::
Db       2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||         ::::::   ||||||   ||||||
Db       2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy        682 Asn 682
              |||
Db       2783 AAT 2785

RESULT 10
US-10-089-678-2
; Sequence 2, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
;  APPLICANT: ASANO, Shinichiro
;  TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN,
;  TITLE OF INVENTION:  NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD
;  FILE REFERENCE: Q68821
;  CURRENT APPLICATION NUMBER: US/10/089,678
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-236140
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: PCT/JP01/06660
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 3504
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1)..(3501)
;   OTHER INFORMATION:
US-10-089-678-2

Alignment Scores:
Pred. No.:              2.07e-95       Length:         3504
Score:                  940.50         Matches:        256
Percent Similarity:     51.8%          Conservative:   111
Best Local Similarity:  36.1%          Mismatches:     288
Query Match:            26.1%          Indels:         56
DB:                     5              Gaps:           23

US-10-782-096-2 (1-682) x US-10-089-678-2 (1-3504)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||      |||:::||||||||||||:::|||||||||            ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              |||:::            |||||||||||||||   |||            |||::::::|||
Db         61 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61
              |||||||||::::::|||   |||   :::         |||   :::      |||:::
Db        121 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||:::   |||                                    |||   |||||||||
Db        181 GAGACGTTTATTAGT--------------------TCATCTACGGTTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                    ||||||   :::|||      ||||||||||||   |||   |||
Db        220 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
                 |||   ::::::|||   ||||||                           :::|||
Db        280 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
              :::|||      ||||||:::||||||      :::   |||         ||||||:::
Db        340 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||      |||:::||||||            :::   :::||||||      |||:::
Db        400 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 459

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178
```

```
Db        460 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 519

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198

Db        520 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 579

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218

Db        580 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 639

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238

Db        640 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 699

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258

Db        700 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 759

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278

Db        760 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 819

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298

Db        820 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 879

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318

Db        880 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 939

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338

Db        940 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTTCTTTT 996

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358

Db        997 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTTGATATACTCAGTGAAATA 1056

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375

Db       1057 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1116

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394

Db       1117 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1176

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414

Db       1177 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1233

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434

Db       1234 CAGGTTCGATCATTA------------------GGGGCGGATTTAGCTAATTACTACGCA 1275

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454

Db       1276 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1335

Qy        455 SerLeu---------ArgTyrGluValProAlaAsn--------------LeuProSer 466

Db       1336 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1395

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486

Db       1396 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1449

Qy        487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504

Db       1450 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTCTAAGAATGCTAGTAGTCCTGCT 1509

Qy        505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524

Db       1510 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1569

Qy        525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541

Db       1570 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1629

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561

Db       1630 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1689

Qy        562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577

Db       1690 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1746

Qy        578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597

Db       1747 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1800

Qy        598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617

Db       1801 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGGAT 1851

Qy        618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636

Db       1852 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 1911

Qy        637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653

Db       1912 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 1971

Qy        654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671

Db       1972 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2031

Qy        672 ArgGluGlnGluValAsnAspLeuPhe 680

Db       2032 GCAAAGAAAGCAGTAAATGGCTTGTTT 2058
```

```
RESULT 11
US-10-089-678-3
; Sequence 3, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
;  APPLICANT: ASANO, Shinichiro
;  TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, A
;  TITLE OF INVENTION:  NOXIOUS ORGANISM-CONTROLLING AGENT AND METHOD
;  FILE REFERENCE: Q68821
;  CURRENT APPLICATION NUMBER: US/10/089,678
;  CURRENT FILING DATE:  2002-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-236140
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: PCT/JP01/06660
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 3690
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-10-089-678-3

Alignment Scores:
Pred. No.:              2.25e-95       Length:         3690
Score:                  940.50         Matches:        256
Percent Similarity:     51.8%          Conservative:   111
Best Local Similarity:  36.1%          Mismatches:     288
Query Match:            26.1%          Indels:         56
DB:                     5              Gaps:           23

US-10-782-096-2 (1-682) x US-10-089-678-3 (1-3690)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
```

```
Db        190 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 246

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db        247 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 306

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61
Db        307 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 366

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        367 GAGACGTTTATTAGT--------------------TCATCTACGGTTCAAACTGGAATT 405

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        406 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 465

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
Db        466 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 525

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        526 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 585

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        586 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 645

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178
Db        646 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 705

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        706 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 765

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        766 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 825

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        826 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 885

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        886 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 945

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        946 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 1005

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db       1006 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 1065

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db       1066 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 1125

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db       1126 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTTCTTTT 1182

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db       1183 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTTGATATACTCAGTGAAATA 1242

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375
Db       1243 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1302

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
Db       1303 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1362

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
Db       1363 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1419

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
Db       1420 CAGGTTCGATCATTA-----------------GGGGCGGATTTAGCTAATTACTACGCA 1461

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
Db       1462 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1521

Qy        455 SerLeu---------ArgTyrGluValProAlaAsn---------------LeuProSer 466
Db       1522 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1581

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
Db       1582 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1635

Qy        487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504
Db       1636 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTTCTAAGAATGCTAGTAGTCCTGCT 1695

Qy        505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524
Db       1696 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1755

Qy        525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541
Db       1756 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1815

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
Db       1816 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1875

Qy        562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577
Db       1876 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1932

Qy        578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597
Db       1933 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1986

Qy        598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617
Db       1987 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGGAT 2037

Qy        618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636
Db       2038 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 2097

Qy        637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653
Db       2098 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 2157

Qy        654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671
Db       2158 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2217

Qy        672 ArgGluGlnGluValAsnAspLeuPhe 680
Db       2218 GCAAAGAAAGCAGTAAATGGCTTGTTT 2244
```

RESULT 12
US-10-032-717-5
; Sequence 5, Application US/10032717
; Publication No. US20020151709A1

```
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Nicholas B. Duck
;  APPLICANT:  Xiang Feng
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Lynn E. Sims
;  TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
;  TITLE OF INVENTION:  Pesticidal Activity Against Coleopterans
;  FILE REFERENCE: 35718/237005
;  CURRENT APPLICATION NUMBER: US/10/032,717
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 60/242,838
;  PRIOR FILING DATE: 2000-10-24
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 2003
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis (truncated)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2001)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: 1218-1
US-10-032-717-5

Alignment Scores:
Pred. No.:               3.72e-94      Length:        2003
Score:                   926.00        Matches:       239
Percent Similarity:      50.8%         Conservative:  116
Best Local Similarity:   34.2%         Mismatches:    268
Query Match:             25.7%         Indels:        76
DB:                      5             Gaps:          23

US-10-782-096-2 (1-682) x US-10-032-717-5 (1-2003)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||      |||:::|||||||||||||:::::::||||||            :::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
                ||||||           ||||||||   |||   :::|||       ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::|||||||:::::|||   |||                   |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||        |||||||||    :::|||    :::||||||             |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||   :::||||||   ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||          :::          :::|||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      ||||||||||||||   :::   |||   :::   |||
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   |||||||||||||||              :::||||||   |||   |||:::
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||                 |||   :::      ::::::
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::    |||   :::   ||||||   |||:::      |||||||||:::
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                 |||   |||||||||   |||||||||||||||||:::||||||:::||||||
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||                  |||   :::   |||   |||
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                    |||      ||||||   |||||||||:::   :::|||      ||||||
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                 :::         ::::::|||:::      ||||||   :::|||:::||||||   :::
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   |||||||||:::      ||||||||||||||||||   |||   |||
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||      :::||||||||||||||||||||||||:::|||
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||         |||               ||||||:::            :::|||
Db        943 AACGTGTCTTCAATTGGTTCC-----------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                 :::|||::::::         |||   |||         ::::::      :::
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 |||||||||         |||::::::      :::|||:::      |||
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
                 ||||||   :::            :::   ||||||   :::         :::|||
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                 ::::::               |||:::|||   |||         ::::::
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer--------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                 |||         |||   :::|||   |||   |||
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||         :::|||:::         |||      |||   :::   :::
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||               |||         |||:::   |||
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT-----------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                         |||||||||   :::         :::||||||
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                 :::|||||||||||||   :::|||::::::         |||         ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                  |||   |||   |||   |||   |||   |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545
```

```
Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::        |||||||||||||||||  |||||||||  |||     |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::  ||||||                ||||||   |||||||||:::|||::::::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::   |||     |||:::                            |||   :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::||||||||||||||                ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                       |||   |||            :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::       |||       |||         |||:::      |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                          :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
                 :::         |||         :::|||:::   |||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

```
RESULT 13
US-10-414-637-5
; Sequence 5, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Nicholas B. Duck
;  APPLICANT:  Xiang Feng
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Lynn E. Sims
;  TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
;  TITLE OF INVENTION:  Pesticidal Activity Against Coleopterans
;  FILE REFERENCE: 35718/237005
;  CURRENT APPLICATION NUMBER: US/10/414,637
;  CURRENT FILING DATE:  2003-04-16
;  PRIOR APPLICATION NUMBER: US/10/032,717
;  PRIOR FILING DATE: 2001-10-23
;  PRIOR APPLICATION NUMBER: 60/242,838
;  PRIOR FILING DATE: 2000-10-24
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 2003
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis (truncated)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2001)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: 1218-1
US-10-414-637-5

Alignment Scores:
Pred. No.:              3.72e-94      Length:         2003
Score:                  926.00        Matches:        239
Percent Similarity:     50.8%         Conservative:   116
Best Local Similarity:  34.2%         Mismatches:     268
Query Match:            25.7%         Indels:         76
DB:                     6             Gaps:           23

US-10-782-096-2 (1-682) x US-10-414-637-5 (1-2003)
```

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||         |||:::||||||||||||||:::::::||||||       ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||         ||||||||   |||     :::|||         ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                          |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
              |||         |||||||||      :::|||   :::||||||              |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||      :::||||||   ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||             :::          :::|||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||   :::   |||   :::        |||
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   |||||||||||||||                :::||||||   |||   |||:::
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||                  |||   :::       ::::::
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::   ||||||   |||:::        |||||||||:::
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                    |||   |||||||||   ||||||||||||||||:::||||||:::||||||
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||                |||   :::   |||   |||
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                       |||      ||||||   ||||||:::     :::|||       ||||||
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                 :::      ::::::|||:::   ||||||   :::|||:::||||||    :::
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   |||||||||:::      ||||||||||||||||||   |||   |||
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||   :::|||||||||||||||||||||||||||:::|||
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||      |||            ||||||:::              :::|||
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
```

```
                ::: |||::::::        |||    |||                 ::::::        ::
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 ||||||||       |||::::::   :::          :::|||:::   |||
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::                :::   ||||||   :::           :::|||
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                  ::::::                    |||:::|||   |||       ::::::
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                    |||              |||      :::|||   |||   |||
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||        :::|||:::      |||      |||       :::   :::
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||          :::|||:::|||                |||      |||:::   |||
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                     |||||||||   :::         :::||||||
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                    :::|||||||||||||   :::|||::::::        |||        ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                |||   |||   |||   |||   |||   |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::         ||||||||||||||||||   |||||||||   |||      |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||                |||||| |||||||||:::|||::::::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::   |||      |||:::                            |||      :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::|||||||||||||||                  ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                    |||   |||            :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::      |||         |||            |||:::         |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                          :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
                 :::         |||      :::|||:::   ||||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

RESULT 14
US-10-032-717-9

```
; Sequence 9, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Nicholas B. Duck
;  APPLICANT:  Xiang Feng
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Lynn E. Sims
;  TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
;  TITLE OF INVENTION:  Pesticidal Activity Against Coleopterans
;  FILE REFERENCE: 35718/237005
;  CURRENT APPLICATION NUMBER: US/10/032,717
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 60/242,838
;  PRIOR FILING DATE: 2000-10-24
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 2010
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2010)
;   OTHER INFORMATION: Maize optimized Cry1218-1
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: mo1218-1
US-10-032-717-9

Alignment Scores:
Pred. No.:              3.74e-94       Length:        2010
Score:                  926.00         Matches:       239
Percent Similarity:     50.8%          Conservative:  116
Best Local Similarity:  34.2%          Mismatches:    268
Query Match:            25.7%          Indels:        76
DB:                     5              Gaps:          23

US-10-782-096-2 (1-682) x US-10-032-717-9 (1-2010)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                 |||         |||:::||||||||||||::::::||||||          ::::::::::::
Db          4 TCCCCCAACAACCAGAACGAGTACGAGATCATCGACGCCACCCCC---TCCACCTCCGTG 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||               |||||||||   |||      :::|||         ::::::|||
Db         61 TCCAACGACTCCAACCGCTACCCCTTCGCCAACGAGCCCACCAACGCCCTCCAGAACATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                          |||:::   |||
Db        121 GACTACAAGGACTACCTCAAGATGTCC------------------GCCGGCAACGCCTCC 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
              |||         |||||||||        :::|||   :::||||||            |||
Db        163 GAGTACCCCGGCTCCCCCGAGGTGCTCGTGTCCGGCCAGGACGCCGCCAAGGCCGCCATC 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||   :::||||||   ||||||||||||
Db        223 GACATCGTGGGCAAGCTCCTCTCCGGCCTCGGCGTGCCCTTCGTGGGCCCCATCGTGTCC 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   ||||||||||          :::          :::|||
Db        283 CTCTACACCCAGCTCATCGACATCCTCTGGCCCTCCGGCGAGAAGTCCCAGTGGGAAATC 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||   :::   |||   :::         |||
Db        343 TTCATGGAGCAGGTGGAGGAGCTCATCAACCAGAAGATCGCCGAGTACGCCCGCAACAAG 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   |||||||||||||||            :::||||||   |||   |||:::
Db        403 GCCCTCTCCGAGCTGGAGGGCCTCGGCAACAACTACCAGCTCTACCTCACCGCCCTGGAG 462
```

```
Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||       |||  |||                |||     :::      ::::::
Db        463 GAGTGGGAGGAGAACCCCAACGGCTCCCGCGCCCTCCGCGACGTGCGCAACCGCTTCGAG 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::  ||||||  |||:::      |||||||||:::
Db        523 ATCCTCGACTCCCTCTTCACCCAGTACATGCCCTCCTTCCGCGTGACCAACTTCGAGGTG 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                    |||   |||||||||   |||||||||||||||||:::||||||:::||||||
Db        583 CCCTTCCTCACCGTGTACGCCATGGCCGCCAACCTCCACCTCCTCCTCCTCAAGGACGCC 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||                |||   :::   |||   |||
Db        643 TCCATCTTCGGCGAGGAGTGGGGCTGGTCCACCACCACCATCAACAACTACTACGACCGC 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                       |||   ||||||   |||||||||:::   :::|||      ||||||
Db        703 CAGATGAAGCTCACCGCCGAGTACTCCGACCACTGCGTGAAGTGGTATGAGACCGGCCTC 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
               :::        ::::::|||:::   ||||||   :::|||:::||||||   :::
Db        763 GCCAAGCTCAAGGGCACCTCCGCCAAGCAGTGGGTGGACTACAACCAGTTCCGCCGCGAG 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              ||||||||   |||||||||:::      ||||||||||||||||||   |||   |||
Db        823 ATGACCCTCGCCGTGCTCGACGTGGTGGCCCTCTTCCCCAACTACGACACCCGCACCTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||      :::||||||||||||||||||||||||:::::|||
Db        883 CCCATGGAGACCAAGGCCCAGCTCACCCGCGAGGTGTACACCGACCCGCTCGGCGCCGTG 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||      |||                |||||:::         :::|||
Db        943 AACGTGTCCTCCATCGGCTCT------------TGGTACGAC---AAGGCCCCAAGCTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                 :::|||::::::      |||   |||          ::::::      :::
Db        988 GGCGTGATCGAGTCCTCCGTGATCCGCCCGCCGCACGTGTTCGACTACATCACCGGCCTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 ||||||||        |||::::::   :::         :::|||:::   |||
Db       1048 ACCGTGTACACCCAGTCCCGCTCCATCTCCTCCGCCCGCTACATCCGCCAC---TGGGCC 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::             :::   |||||   :::        :::|||
Db       1105 GGCCACCAGATCTCCTACCACCGCGTGTCCCGCGGCTCCAACCTCCAGCAGATGTACGGC 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                       ::::::                ||||::|||   |||      ::::::
Db       1165 ACCAACCAGAACCTCCACTCCACCTCCACCTTCGACTTCACCAACTACGACATCTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                 |||                  |||   :::|||   |||   |||
Db       1225 ACCCTCTCCAAGGACGCCGTGCTCTCGACATCGTGTACCCCGGCTACACC--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||      :::|||:::      |||      |||      :::   :::
Db       1276 ---TACATCTTC------TTCGGCATGCCGGAGGTGGAGTTCTTCATGGTGAACCAGCTC 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||               |||      |||:::   |||
Db       1327 AACAACACCCGCAAGACCCTCAAATAC------------AACCCCGTGTCCAAGGACATC 1374

Qy        470 LeuSer------------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                        |||||||||   :::         :::||||||
Db       1375 ATCGCCTCCACCCGCGACTCCGAGCTCGAGCTCCCCCCCGAGACCTCCGACCAGCCCAAC 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                       :::||||||||||||   :::|||::::::         |||      ::::::
Db       1435 TACGAGTCCTACTCCCACCGCCTCTGCCACATCACCTCCATCCCCGCCACCGGCAACACC 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                   |||   |||   |||   |||   |||   |||
Db       1495 ACCGGCCTCGTGCCG---------GTGTTCTCCTGGACCCACCGCTCTGCAGACCTCAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::   ||||||||||||||||||   |||||||||   |||      |||
Db       1546 AACACCATCTACTCCGACAAGATCACCCAGATCCCCGCCGTGAAGTGCTGG---GACAAC 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||               ||||||   |||||||||:::|||::::::
Db       1603 CTCCCCTTCGTGCCCGTGGTGAAGGGCCCCGGCCACACCGGCGGCGACCTCCTCCAGTAC 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::   |||      |||:::                              |||   :::
Db       1663 AACCGCTCCACCGGCTCCGTGGGCACCCTCTTCCTCGCCCGCTACGGCCTCGCCCTGGAG 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::||||||||||||||                  ||||||
Db       1723 AAGGCCGGCAAGTACCGCGTGCGCCTCCGCTACGCCACT---------------GACGCC 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                 |||   |||         :::   |||
Db       1768 GACATCGTGCTCCACGTGAACGACGCCCAGATCCAG---------------ATGCCCAAG 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::      |||         |||         |||:::      |||
Db       1813 ACCATGAACCCCGGCGAGGACCTCACCTCCAAGACCTTCAAGGTGGCCGACGCCATCACC 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg------------------- 644
                       :::   :::      :::      :::
Db       1873 ACCCTCAACCTCGCCACCGACTCCTCCCTCGCCCTCAAGCACAACCTCGGCGAGGACCCC 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       1933 AACTCCACCCTCTCCGGCATCGTGTACGTGGACCGCATCGAGTTCATCCCCGTGGAC 1989
```

RESULT 15
US-10-032-717-15
; Sequence 15, Application US/10032717
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 2010
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis (truncated)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2010)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: 1218-1A
US-10-032-717-15

Alignment Scores:

Pred. No.:	3.74e-94	Length:	2010
Score:	926.00	Matches:	239
Percent Similarity:	50.8%	Conservative:	116
Best Local Similarity:	34.2%	Mismatches:	268
Query Match:	25.7%	Indels:	76

```
Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
```

```
DB:                    5            Gaps:         23
US-10-782-096-2 (1-682) x US-10-032-717-15 (1-2010)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db        4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db       61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db      283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db      343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db      403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db      463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db      523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db      583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db      643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db      703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db      763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db      823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db      883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db      943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db      988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db     1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db     1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db     1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy      416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db     1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db     1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db     1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy      470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db     1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db     1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db     1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db     1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db     1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy      559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db     1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db     1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy      592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db     1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy      612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db     1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy      632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db     1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy      645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db     1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

Search completed: January 21, 2006, 04:23:59

Job time : 1474 secs

```
                    GenCore version 5.1.7
         Copyright (c) 1993 - 2006  Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:        January 20, 2006, 23:21:18 ; Search time 296 Seconds
                                 (without alignments)
                                 1899.170 Million cell updates/sec

Title:         US-10-782-096-2
Perfect score: 3597
Sequence:      1 MNSYKNKNEYEMLDALRINS..........TFPNQSLEKREQEVNDLFIN 682

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

Searched:      6049916 seqs, 412136615 residues

Total number of hits satisfying chosen parameters:      12099832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10782096/runat_20012006_095216_24440/app_query.fasta_1.839
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10782096_@CGN_1_1_121_@runat_20012006_095216_24440
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :     Published_Applications_NA_New:*
               1:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
               2:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
               3:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
               4:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
               5:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
               6:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
               7:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
               8:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
               9:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
               10:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
               11:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

  Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	948	26.4	3621	8	US-11-058-727-1	Sequence 1, Appli
2	948	26.4	3621	8	US-11-108-389-1	Sequence 1, Appli
3	948	26.4	4874	8	US-11-058-727-17	Sequence 17, Appl
4	948	26.4	4874	8	US-11-108-389-17	Sequence 17, Appl
5	926	25.7	2010	8	US-11-058-727-5	Sequence 5, Appli
6	926	25.7	2010	8	US-11-058-727-11	Sequence 11, Appl
7	926	25.7	2010	8	US-11-108-389-5	Sequence 5, Appli
8	926	25.7	2010	8	US-11-108-389-11	Sequence 11, Appl
9	925.5	25.7	2022	8	US-11-058-727-49	Sequence 49, Appl
10	925.5	25.7	2022	8	US-11-058-727-81	Sequence 81, Appl
11	925.5	25.7	2022	8	US-11-108-389-49	Sequence 49, Appl
12	925.5	25.7	2022	8	US-11-108-389-81	Sequence 81, Appl
13	925	25.7	2022	8	US-11-058-727-7	Sequence 7, Appli
14	925	25.7	2022	8	US-11-058-727-25	Sequence 25, Appl
15	925	25.7	2022	8	US-11-058-727-29	Sequence 29, Appl
16	925	25.7	2022	8	US-11-058-727-33	Sequence 33, Appl
17	925	25.7	2022	8	US-11-058-727-69	Sequence 69, Appl
18	925	25.7	2022	8	US-11-108-389-7	Sequence 7, Appli
19	925	25.7	2022	8	US-11-108-389-25	Sequence 25, Appl
20	925	25.7	2022	8	US-11-108-389-29	Sequence 29, Appl
21	925	25.7	2022	8	US-11-108-389-33	Sequence 33, Appl
22	925	25.7	2022	8	US-11-108-389-69	Sequence 69, Appl
23	924.5	25.7	2022	8	US-11-058-727-43	Sequence 43, Appl
24	924.5	25.7	2022	8	US-11-058-727-75	Sequence 75, Appl
25	924.5	25.7	2022	8	US-11-108-389-43	Sequence 43, Appl
26	924.5	25.7	2022	8	US-11-108-389-75	Sequence 75, Appl
27	924	25.7	2025	8	US-11-058-727-45	Sequence 45, Appl
28	924	25.7	2025	8	US-11-058-727-77	Sequence 77, Appl
29	924	25.7	2025	8	US-11-108-389-45	Sequence 45, Appl
30	924	25.7	2025	8	US-11-108-389-77	Sequence 77, Appl
31	924	25.7	3633	8	US-11-058-727-3	Sequence 3, Appli
32	924	25.7	3633	8	US-11-108-389-3	Sequence 3, Appli
33	924	25.7	6613	8	US-11-058-727-18	Sequence 18, Appl
34	924	25.7	6613	8	US-11-108-389-18	Sequence 18, Appl
35	920	25.6	2025	8	US-11-058-727-47	Sequence 47, Appl
36	920	25.6	2025	8	US-11-058-727-79	Sequence 79, Appl
37	920	25.6	2025	8	US-11-108-389-47	Sequence 47, Appl
38	920	25.6	2025	8	US-11-108-389-79	Sequence 79, Appl
39	916	25.5	2022	8	US-11-058-727-21	Sequence 21, Appl
40	916	25.5	2022	8	US-11-058-727-63	Sequence 63, Appl
41	916	25.5	2022	8	US-11-058-727-65	Sequence 65, Appl
42	916	25.5	2022	8	US-11-058-727-67	Sequence 67, Appl
43	916	25.5	2022	8	US-11-108-389-21	Sequence 21, Appl
44	916	25.5	2022	8	US-11-108-389-63	Sequence 63, Appl
45	916	25.5	2022	8	US-11-108-389-65	Sequence 65, Appl

ALIGNMENTS

```
RESULT 1
US-11-058-727-1
; Sequence 1, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Rafael Herrmann
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Albert L. Lu
;  APPLICANT:  Billy Fred McCutchen
;  APPLICANT:  James K. Presnail
;  APPLICANT:  James F.H. Wong
;  APPLICANT:  Cao-Guo Yu
;  TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
;  TITLE OF INVENTION:  Activity
;  FILE REFERENCE: 35718/287809
;  CURRENT APPLICATION NUMBER: US/11/058,727
;  CURRENT FILING DATE:  2005-02-15
;  PRIOR APPLICATION NUMBER: 60/391,786
;  PRIOR FILING DATE: 2002-06-26
;  PRIOR APPLICATION NUMBER: 60/460,787
;  PRIOR FILING DATE: 2003-04-04
;  PRIOR APPLICATION NUMBER: 10/606,320
;  PRIOR FILING DATE: 2003-06-25
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
```

```
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)...(3621)
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (0)...(0)
;    OTHER INFORMATION: Cry1218-1
US-11-058-727-1

Alignment Scores:
Pred. No.:               4.05e-102     Length:         3621
Score:                   948.00        Matches:        248
Percent Similarity:      50.9%         Conservative:   119
Best Local Similarity:   34.4%         Mismatches:     276
Query Match:             26.4%         Indels:         78
DB:                      8             Gaps:           25

US-10-782-096-2 (1-682) x US-11-058-727-1 (1-3621)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||      |||:::|||||||||||||:::::::||||||            :::::::::::
Db        4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||         |||||||||    |||     :::|||         ::::::|||
Db       61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
            :::|||||||:::::|||   |||                           |||:::    |||
Db      121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
               |||         |||||||||        :::|||    :::||||||             |||
Db      163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
            :::    |||||     :::||||||     ||||||||||||
Db      223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
            :::     ::::::||||||   |||||||||                :::           :::|||
Db      283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
               |||          ||||||||||||       :::     :::        |||
Db      343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
            ||||||      |||||||||||||||                :::||||||    |||    |||:::
Db      403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTCAA 462

Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
               |||         |||    |||                |||      :::      ::::::
Db      463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
            |||:::|||:::      |||     :::     ||||||    |||:::          |||||||||:::
Db      523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                  |||      ||||||||      ||||||||||||||||||:::||||||:::||||||
Db      583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                  ||||||      :::|||                   |||     :::     |||
Db      643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                  |||         ||||||   |||||||||:::      :::|||        ||||||
Db      703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
               :::         ::::::|||:::    ||||||    :::|||:::|||||          :::
Db      763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
            |||||||||     |||||||||:::         |||||||||||||||||      |||    |||
Db      823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
            |||:::        |||     :::||||||||||||||||||||||||||||:::|||
Db      883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
               |||        |||                      ||||||:::               :::|||
Db      943 AACGTGTCTTCAATTGGTTCC-----------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                  :::|||::::::      |||     |||               ::::::          :::
Db      988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
               ||||||||        |||::::::     :::             :::|||:::      |||
Db     1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
            |||||     :::                    :::     ||||||     :::          :::|||
Db     1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                 ::::::                        |||:::|||     |||          ::::::
Db     1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy      416 IleGluSer--------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
               |||                      |||      :::|||    |||    |||
Db     1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
               ||||||        :::|||:::          |||          |||          :::    :::
Db     1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
            ||||||         :::|||:::|||                |||         |||:::    |||
Db     1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy      470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
            ::::::                          |||||||||     :::          :::||||||
Db     1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                  :::|||||||||||||     :::|||::::::          |||          ::::::
Db     1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
            :::|||     :::                   |||    |||    |||    |||    |||    |||
Db     1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
            |||     :::          |||||||||||||||     |||||||||    |||          |||
Db     1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
            :::     ||||||                        ||||||     |||||||||:::|||::::::
Db     1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy      559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
            :::     |||      |||:::
Db     1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
            :::            ||||||:::||||||||||||||
Db     1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy      592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                  |||     |||               :::    |||
Db     1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812
```

```
Qy       612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db      1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy       632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644
Db      1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy       645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db      1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy       663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db      1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy       682 Asn 682
Db      2053 AAT 2055
```

RESULT 2
US-11-108-389-1
; Sequence 1, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3621)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Cry1218-1
US-11-108-389-1

Alignment Scores:

Pred. No.:	4.05e-102	Length:	3621
Score:	948.00	Matches:	248
Percent Similarity:	50.9%	Conservative:	119
Best Local Similarity:	34.4%	Mismatches:	276
Query Match:	26.4%	Indels:	78
DB:	8	Gaps:	25

US-10-782-096-2 (1-682) x US-11-108-389-1 (1-3621)

```
Qy         3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db         4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy        23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db        61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy        43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db       121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy        63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db       163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy        82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db       223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy       102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db       283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy       120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db       343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy       140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db       403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy       160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db       463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy       179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db       523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy       199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db       583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy       219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db       643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy       239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db       703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy       259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db       763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy       279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db       823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy       299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db       883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy       319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db       943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy       339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db       988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy       359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db      1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104
```

```
Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||              :::   ||||||   :::        :::|||
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                       ::::::                 |||:::|||   |||      ::::::
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                    |||               |||      :::|||   |||   |||
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||         :::|||:::      |||      |||      :::   :::
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||               |||      |||:::   |||
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                    |||||||||   :::         :::||||||
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                       :::||||||||||||   :::|||::::::         |||      ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::               |||   |||   |||   |||   |||   |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::      |||||||||||||||||   |||||||||   |||      |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::      ||||||            ||||||   |||||||||:::|||::::::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
              :::   |||      |||:::                           |||      :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::||||||||||||||||                  ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                 |||   |||            :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::         |||      |||            |||:::   |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                       :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||         ::::::   ||||||   ||||||
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
              |||
Db       2053 AAT 2055
```

```
RESULT 3
US-11-058-727-17
; Sequence 17, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Rafael Herrmann
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Albert L. Lu
;  APPLICANT:  Billy Fred McCutchen
;  APPLICANT:  James K. Presnail
;  APPLICANT:  James F.H. Wong
;  APPLICANT:  Cao-Guo Yu
;  TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
;  TITLE OF INVENTION:  Activity
;  FILE REFERENCE: 35718/287809
;  CURRENT APPLICATION NUMBER: US/11/058,727
;  CURRENT FILING DATE:  2005-02-15
;  PRIOR APPLICATION NUMBER: 60/391,786
;  PRIOR FILING DATE: 2002-06-26
;  PRIOR APPLICATION NUMBER: 60/460,787
;  PRIOR FILING DATE: 2003-04-04
;  PRIOR APPLICATION NUMBER: 10/606,320
;  PRIOR FILING DATE: 2003-06-25
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 4874
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Genomic DNA 1218-1
US-11-058-727-17

Alignment Scores:
Pred. No.:               6.76e-102     Length:        4874
Score:                   948.00        Matches:       248
Percent Similarity:      50.9%         Conservative:  119
Best Local Similarity:   34.4%         Mismatches:    276
Query Match:             26.4%         Indels:        78
DB:                      8             Gaps:          25

US-10-782-096-2 (1-682) x US-11-058-727-17 (1-4874)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||      |||:::||||||||||||::::::||||||         ::::::::::::
Db        734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||         |||||||||   |||      :::|||         ::::::|||
Db        791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                        |||:::   |||
Db        851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
              |||         |||||||||      :::|||   :::||||||            |||
Db        893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||   :::||||||   ||||||||||||
Db        953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   ||||||||||         :::         :::|||
Db       1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||   :::   |||   :::      |||
```

```
Db       1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db       1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db       1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db       1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db       1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db       1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db       1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db       1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db       1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db       1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db       1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db       1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db       2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db       2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db       2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db       2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db       2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db       2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy        682 Asn 682
Db       2783 AAT 2785
```

RESULT 4
US-11-108-389-17
; Sequence 17, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320

```
;  PRIOR FILING DATE: 2003-06-25
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 4874
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Genomic DNA 1218-1
US-11-108-389-17

Alignment Scores:
Pred. No.:               6.76e-102      Length:         4874
Score:                   948.00         Matches:        248
Percent Similarity:      50.9%          Conservative:   119
Best Local Similarity:   34.4%          Mismatches:     276
Query Match:             26.4%          Indels:         78
DB:                      8              Gaps:           25

US-10-782-096-2 (1-682) x US-11-108-389-17 (1-4874)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db      734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790
Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db      791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850
Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892
Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952
Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012
Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db     1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072
Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db     1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132
Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db     1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192
Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db     1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252
Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db     1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312
Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db     1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372
Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db     1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432
Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db     1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492
Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db     1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552
Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db     1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612
Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db     1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672
Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db     1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717
Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db     1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777
Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db     1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834
Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db     1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894
Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db     1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954
Qy      416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db     1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005
Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db     2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056
Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db     2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104
Qy      470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db     2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164
Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db     2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224
Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db     2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275
Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db     2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332
Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db     2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392
Qy      559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db     2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452
Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db     2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497
```

```
Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db       2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db       2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db       2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy        682 Asn 682
Db       2783 AAT 2785
```

RESULT 5
US-11-058-727-5
; Sequence 5, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 2010
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2010)
;   OTHER INFORMATION: Maize optimized Cry1218-1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: mo1218-1
US-11-058-727-5

Alignment Scores:

Pred. No.:	6.54e-100	Length:	2010
Score:	926.00	Matches:	239
Percent Similarity:	50.8%	Conservative:	116
Best Local Similarity:	34.2%	Mismatches:	268
Query Match:	25.7%	Indels:	76
DB:	8	Gaps:	23

US-10-782-096-2 (1-682) x US-11-058-727-5 (1-2010)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 TCCCCCAACAACCAGAACGAGTACGAGATCATCGACGCCACCCCC---TCCACCTCCGTG 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAACGACTCCAACCGCTACCCCTTCGCCAACGAGCCCACCAACGCCCTCCAGAACATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GACTACAAGGACTACCTCAAGATGTCC-----------------GCCGGCAACGCCTCC 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAGTACCCCGGCTCCCCCGAGGTGCTCGTGTCCGGCCAGGACGCCGCCAAGGCCGCCATC 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GACATCGTGGGCAAGCTCCTCTCCGGCCTCGGCGTGCCCTTCGTGGGCCCCATCGTGTCC 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTCTACACCCAGCTCATCGACATCCTCTGGCCCTCCGGCGAGAAGTCCCAGTGGGAAATC 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTCATGGAGCAGGTGGAGGAGCTCATCAACCAGAAGATCGCCGAGTACGCCCGCAACAAG 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCCCTCTCCGAGCTGGAGGGCCTCGGCAACAACTACCAGCTCTACCTCACCGCCCTGGAG 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAGTGGGAGGAGAACCCCAACGGCTCCCGCGCCCTCCGCGACGTGCGCAACCGCTTCGAG 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTCGACTCCCTCTTCACCCAGTACATGCCCTCCTTCCGCGTGACCAACTTCGAGGTG 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCCTTCCTCACCGTGTACGCCATGGCCGCCAACCTCCACCTCCTCCTCCTCAAGGACGCC 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCCATCTTCGGCGAGGAGTGGGGCTGGTCCACCACCACCATCAACAACTACTACGACCGC 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAGATGAAGCTCACCGCCGAGTACTCCGACCACTGCGTGAAGTGGTATGAGACCGGCCTC 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCCAAGCTCAAGGGCACCTCCGCCAAGCAGTGGGTGGACTACAACCAGTTCCGCCGCGAG 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACCCTCGCCGTGCTCGACGTGGTGGCCCTCTTCCCCAACTACGACACCCGCACCTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCCATGGAGACCAAGGCCCAGCTCACCCGCGAGGTGTACACCGACCCGCTCGGCGCCGTG 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCCTCCATCGGCTCT------------TGGTACGAC---AAGGCCCCAAGCTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
```

```
Db        988 GGCGTGATCGAGTCCTCCGTGATCCGCCCGCCGCACGTGTTCGACTACATCACCGGCCTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 ||||||||      |||::::::   :::        :::|||:::   |||
Db       1048 ACCGTGTACACCCAGTCCCGCTCCATCTCCTCCGCCCGCTACATCCGCCAC---TGGGCC 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::                   :::  ||||||   :::        :::|||
Db       1105 GGCCACCAGATCTCCTACCACCGCGTGTCCCGCGGCTCCAACCTCCAGCAGATGTACGGC 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                       ::::::                    |||:::|||   |||      ::::::
Db       1165 ACCAACCAGAACCTCCACTCCACCTCCACCTTCGACTTCACCAACTACGACATCTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                    |||                 |||      :::|||   |||   |||
Db       1225 ACCCTCTCCAAGGACGCCGTGCTCCTCGACATCGTGTACCCCGGCTACACC--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||      :::|||:::      |||      |||      :::   :::
Db       1276 ---TACATCTTC------TTCGGCATGCCGGAGGTGGAGTTCTTCATGGTGAACCAGCTC 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||                |||      |||:::   |||
Db       1327 AACAACACCCGCAAGACCCTCAAATAC------------AACCCCGTGTCCAAGGACATC 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                    |||||||||   :::         :::||||||
Db       1375 ATCGCCTCCACCCGCGACTCCGAGCTCGAGCTCCCCCCCGAGACCTCCGACCAGCCCAAC 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                    :::||||||||||||      :::|||::::::      |||         ::::::
Db       1435 TACGAGTCCTACTCCCACCGCCTCTGCCACATCACCTCCATCCCCGCCACCGGCAACACC 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                |||   |||   |||   |||   |||   |||
Db       1495 ACCGGCCTCGTGCCG---------GTGTTCTCCTGGACCCACCGCTCTGCAGACCTCAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::      ||||||||||||||||||   |||||||||   |||      |||
Db       1546 AACACCATCTACTCCGACAAGATCACCCAGATCCCCGCCGTGAAGTGCTGG---GACAAC 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||                ||||||   |||||||||:::|||::::::
Db       1603 CTCCCCTTCGTGCCCGTGGTGAAGGGCCCCGGCCACACCGGCGGCGACCTCCTCCAGTAC 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
              :::   |||      |||:::                            |||      :::
Db       1663 AACCGCTCCACCGGCTCCGTGGGCACCCTCTTCCTCGCCCGCTACGGCCTCGCCCTGGAG 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::|||||||||||||||                  ||||||
Db       1723 AAGGCCGGCAAGTACCGCGTGCGCCTCCGCTACGCCACT---------------GACGCC 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                    |||   |||         :::   |||
Db       1768 GACATCGTGCTCCACGTGAACGACGCCCAGATCCAG---------------ATGCCCAAG 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::         |||      |||               |||:::         |||
Db       1813 ACCATGAACCCCGGCGAGGACCTCACCTCCAAGACCTTCAAGGTGGCCGACGCCATCACC 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                          :::   :::         :::
Db       1873 ACCCTCAACCTCGCCACCGACTCCTCCCTCGCCCTCAAGCACAACCTCGGCGAGGACCCC 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       1933 AACTCCACCCTCTCCGGCATCGTGTACGTGGACCGCATCGAGTTCATCCCCGTGGAC 1989
```

RESULT 6
US-11-058-727-11
; Sequence 11, Application US/11058727

```
; Publication No. US20050261483A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Rafael Herrmann
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Albert L. Lu
;  APPLICANT:  Billy Fred McCutchen
;  APPLICANT:  James K. Presnail
;  APPLICANT:  James F.H. Wong
;  APPLICANT:  Cao-Guo Yu
;  TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
;  TITLE OF INVENTION:  Activity
;  FILE REFERENCE: 35718/287809
;  CURRENT APPLICATION NUMBER: US/11/058,727
;  CURRENT FILING DATE:  2005-02-15
;  PRIOR APPLICATION NUMBER: 60/391,786
;  PRIOR FILING DATE: 2002-06-26
;  PRIOR APPLICATION NUMBER: 60/460,787
;  PRIOR FILING DATE: 2003-04-04
;  PRIOR APPLICATION NUMBER: 10/606,320
;  PRIOR FILING DATE: 2003-06-25
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 2010
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis (truncated)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2010)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: 1218-1A
US-11-058-727-11

Alignment Scores:
Pred. No.:              6.54e-100      Length:        2010
Score:                  926.00         Matches:       239
Percent Similarity:     50.8%          Conservative:  116
Best Local Similarity:  34.2%          Mismatches:    268
Query Match:            25.7%          Indels:        76
DB:                     8              Gaps:          23

US-10-782-096-2 (1-682) x US-11-058-727-11 (1-2010)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||         |||:::||||||||||||::::::||||||            ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
                ||||||            |||||||||   |||   :::|||         ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                               |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||         :::|||   :::||||||            |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                :::   ||||||   :::||||||   ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
                :::   ::::::||||||   |||||||||         :::         :::|||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||   :::   |||   :::      |||
```

```
Db     343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy     140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159

Db     403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy     160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178

Db     463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy     179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198

Db     523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy     199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218

Db     583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy     219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238

Db     643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy     239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258

Db     703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy     259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278

Db     763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy     279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298

Db     823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy     299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318

Db     883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy     319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338

Db     943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy     339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358

Db     988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy     359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376

Db    1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy     377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395

Db    1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy     396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415

Db    1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy     416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430

Db    1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy     431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449

Db    1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy     450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469

Db    1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy     470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482

Db    1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy     483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502

Db    1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy     503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522

Db    1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy     523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542

Db    1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy     543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558

Db    1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy     559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571

Db    1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy     572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591

Db    1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy     592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611

Db    1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy     612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631

Db    1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy     632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644

Db    1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy     645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662

Db    1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

RESULT 7
US-11-108-389-5
; Sequence 5, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```
;   NAME/KEY: CDS
;   LOCATION: (1)...(2010)
;   OTHER INFORMATION: Maize optimized Cry1218-1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: mo1218-1
US-11-108-389-5

Alignment Scores:
Pred. No.:               6.54e-100      Length:         2010
Score:                   926.00         Matches:        239
Percent Similarity:      50.8%          Conservative:   116
Best Local Similarity:   34.2%          Mismatches:     268
Query Match:             25.7%          Indels:         76
DB:                      8              Gaps:           23

US-10-782-096-2 (1-682) x US-11-108-389-5 (1-2010)

Qy      3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db      4 TCCCCCAACAACCAGAACGAGTACGAGATCATCGACGCCACCCCC---TCCACCTCCGTG 60

Qy     23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db     61 TCCAACGACTCCAACCGCTACCCCTTCGCCAACGAGCCCACCAACGCCCTCCAGAACATG 120

Qy     43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db    121 GACTACAAGGACTACCTCAAGATGTCC-----------------GCCGGCAACGCCTCC 162

Qy     63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db    163 GAGTACCCCGGCTCCCCCGAGGTGCTCGTGTCCGGCCAGGACGCCGCCAAGGCCGCCATC 222

Qy     82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db    223 GACATCGTGGGCAAGCTCCTCTCCGGCCTCGGCGTGCCCTTCGTGGGCCCCATCGTGTCC 282

Qy    102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db    283 CTCTACACCCAGCTCATCGACATCCTCTGGCCCTCCGGCGAGAAGTCCCAGTGGGAAATC 342

Qy    120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db    343 TTCATGGAGCAGGTGGAGGAGCTCATCAACCAGAAGATCGCCGAGTACGCCCGCAACAAG 402

Qy    140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db    403 GCCCTCTCCGAGCTGGAGGGCCTCGGCAACAACTACCAGCTCTACCTCACCGCCCTGGAG 462

Qy    160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db    463 GAGTGGGAGGAGAACCCCAACGGCTCCCGCGCCCTCCGCGACGTGCGCAACCGCTTCGAG 522

Qy    179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db    523 ATCCTCGACTCCCTCTTCACCCAGTACATGCCCTCCTTCCGCGTGACCAACTTCGAGGTG 582

Qy    199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db    583 CCCTTCCTCACCGTGTACGCCATGGCCGCCAACCTCCACCTCCTCCTCCTCAAGGACGCC 642

Qy    219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db    643 TCCATCTTCGGCGAGGAGTGGGGCTGGTCCACCACCACCATCAACAACTACTACGACCGC 702

Qy    239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db    703 CAGATGAAGCTCACCGCCGAGTACTCCGACCACTGCGTGAAGTGGTATGAGACCGGCCTC 762

Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db    763 GCCAAGCTCAAGGGCACCTCCGCCAAGCAGTGGGTGGACTACAACCAGTTCCGCCGCGAG 822

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db    823 ATGACCCTCGCCGTGCTCGACGTGGTGGCCCTCTTCCCCAACTACGACACCCGCACCTAC 882

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db    883 CCCATGGAGACCAAGGCCCAGCTCACCCGCGAGGTGTACACCGACCCGCTCGGCGCCGTG 942

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db    943 AACGTGTCCTCCATCGGCTCT------------TGGTACGAC---AAGGCCCCAAGCTTC 987

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db    988 GGCGTGATCGAGTCCTCCGTGATCCGCCCGCCGCACGTGTTCGACTACATCACCGGCCTC 1047

Qy    359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db   1048 ACCGTGTACACCCAGTCCCGCTCCATCTCCTCCGCCCGCTACATCCGCCAC---TGGGCC 1104

Qy    377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db   1105 GGCCACCAGATCTCCTACCACCGCGTGTCCCGCGGCTCCAACCTCCAGCAGATGTACGGC 1164

Qy    396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db   1165 ACCAACCAGAACCTCCACTCCACCTCCACCTTCGACTTCACCAACTACGACATCTACAAG 1224

Qy    416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db   1225 ACCCTCTCCAAGGACGCCGTGCTCCTCGACATCGTGTACCCCGGCTACACC--------- 1275

Qy    431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db   1276 ---TACATCTTC------TTCGGCATGCCGGAGGTGGAGTTCTTCATGGTGAACCAGCTC 1326

Qy    450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db   1327 AACAACACCCGCAAGACCCTCAAATAC------------AACCCCGTGTCCAAGGACATC 1374

Qy    470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db   1375 ATCGCCTCCACCCGCGACTCCGAGCTCGAGCTCCCCCCCGAGACCTCCGACCAGCCCAAC 1434

Qy    483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db   1435 TACGAGTCCTACTCCCACCGCCTCTGCCACATCACCTCCATCCCCGCCACCGGCAACACC 1494

Qy    503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db   1495 ACCGGCCTCGTGCCG---------GTGTTCTCCTGGACCCACCGCTCTGCAGACCTCAAC 1545

Qy    523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db   1546 AACACCATCTACTCCGACAAGATCACCCAGATCCCCGCCGTGAAGTGCTGG---GACAAC 1602

Qy    543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db   1603 CTCCCCTTCGTGCCCGTGGTGAAGGGCCCCGGCCACACCGGCGGCGACCTCCTCCAGTAC 1662

Qy    559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
Db   1663 AACCGCTCCACCGGCTCCGTGGGCACCCTCTTCCTCGCCCGCTACGGCCTCGCCCTGGAG 1722

Qy    572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db   1723 AAGGCCGGCAAGTACCGCGTGCGCCTCCGCTACGCCACT---------------GACGCC 1767

Qy    592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db   1768 GACATCGTGCTCCACGTGAACGACGCCCAGATCCAG---------------ATGCCCAAG 1812
```

```
Qy     612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db    1813 ACCATGAACCCCGGCGAGGACCTCACCTCCAAGACCTTCAAGGTGGCCGACGCCATCACC 1872

Qy     632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644
Db    1873 ACCCTCAACCTCGCCACCGACTCCTCCCTCGCCCTCAAGCACAACCTCGGCGAGGACCCC 1932

Qy     645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db    1933 AACTCCACCCTCTCCGGCATCGTGTACGTGGACCGCATCGAGTTCATCCCCGTGGAC 1989
```

RESULT 8
US-11-108-389-11
; Sequence 11, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2010)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: 1218-1A
US-11-108-389-11

Alignment Scores:

Pred. No.:	6.54e-100	Length:	2010
Score:	926.00	Matches:	239
Percent Similarity:	50.8%	Conservative:	116
Best Local Similarity:	34.2%	Mismatches:	268
Query Match:	25.7%	Indels:	76
DB:	8	Gaps:	23

US-10-782-096-2 (1-682) x US-11-108-389-11 (1-2010)

```
Qy       3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db       4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy      23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db      61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy      43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db     121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy      63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db     163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy      82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db     223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy     102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db     283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy     120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db     343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy     140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db     403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy     160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db     463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy     179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db     523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy     199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db     583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy     219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db     643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy     239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db     703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy     259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db     763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy     279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db     823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy     299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db     883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy     319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db     943 AACGTGTCTTCAATTGGTTCC-----------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy     339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db     988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy     359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db    1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy     377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db    1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy     396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db    1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224
```

```
Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

RESULT 9
US-11-058-727-49
; Sequence 49, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2022)
US-11-058-727-49

Alignment Scores:

Pred. No.:	7.6e-100	Length:	2022
Score:	925.50	Matches:	243
Percent Similarity:	51.3%	Conservative:	119
Best Local Similarity:	34.4%	Mismatches:	259
Query Match:	25.7%	Indels:	85
DB:	8	Gaps:	26

US-10-782-096-2 (1-682) x US-11-058-727-49 (1-2022)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAsp---------AsnArgArgAlaLeu--------- 173
Db        463 GAATGGGAAGAAAATCCAAATGGTTCCCGGTTTCGAAGTCGACAAGCCTTACGAGATGTG 522

Qy        174 ValThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPheLysGlu 193
Db        523 CGAAATCGATTTGAAATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTG 582

Qy        194 ArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIle 213
Db        583 ACAAATTTTGAAGTACCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTG 642
```

```
Qy    214 LeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArg 233
Db    643 TTATTAAAGGACGCGTCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAAT 702

Qy    234 AspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPhe 253
Db    703 AACTATTATGATCGTCAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGG 762

Qy    254 TyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsn 273
Db    763 TATGAAACTGGTTTAGCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAAC 822

Qy    274 ArgPheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyr 293
Db    823 CAATTCCGTAGAGAAATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTAT 882

Qy    294 AspProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAsp 313
Db    883 GACACACGCACGTACCCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGAT 942

Qy    314 ProValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnPro 333
Db    943 CCACTGGGCGCGGTAAACGTGTCTTCAATTGGTTCC------------TGGTATGAC--- 987

Qy    334 AsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThr 353
Db    988 AAAGCACCTTCTTTCGGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGAT 1047

Qy    354 TrpLeuAsnArgIlePheValTyrThr------ArgThrLeuGlyAsnMetSerAspVal 371
Db   1048 TATATAACGGGACTCACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATA 1107

Qy    372 ArgAsnIleTrpGlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIle 390
Db   1108 AGACAT---TGGGCTGGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTT 1164

Qy    391 ThrHisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsn 410
Db   1165 CAACAAATGTATGGAACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAAT 1224

Qy    411 LeuSerValPheSerIleGluSer---------------LeuAlaArgIleTyrLeuGly 425
Db   1225 TATGATATTTACAAGACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGT 1284

Qy    426 GlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePhe 444
Db   1285 TATACG------------TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTC 1326

Qy    445 AsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeu 464
Db   1327 ATGGTAAACCAATTGAATAATACCAGAAAGACGTTAAAGTAT------------AATCCA 1374

Qy    465 ProSerGlnThrIleLeuSer---------------------GluLeuProGlyLysAsp 477
Db   1375 GTTTCCAAAGATATTATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACT 1434

Qy    478 LysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAsp 497
Db   1435 TCAGATCAACCAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCC 1494

Qy    498 AlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThr 517
Db   1495 GCGACGGGTAACACTACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGA 1545

Qy    518 SerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLys 537
Db   1546 AGTGCAGATTTAAACAATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAA 1605

Qy    538 GlyTrpGlyGlyAsnIleGlyPheValIle------------ProGlyProThrGlyGly 553
Db   1606 TGTTGG---GATAATTTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGG 1662

Qy    554 AsnLeuValLysValSerAspSerTrpHisSerLeuLysVal------------------ 567
Db   1663 GATTTATTACAGTATAATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATAT 1722

Qy    568 ---GlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuVal 586
Db   1723 GGCCTAGCATTAGAAAAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT------ 1776

Qy    587 ThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePhe 606
Db   1777 ---------GATGCAGATATTGTATTGCATGTAAACGATGCTCAGATTCAG--------- 1818

Qy    607 AspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyr 626
Db   1819 ------ATGCCAAAAACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTT 1872

Qy    627 IleAspValProGlyIlePheThrProSerIle------------AsnProLeuIleArg 642
Db   1873 GCAGAT------GCTATCACAACAGTTAATTTAGCAACAGATAGTTCGGTTGCAGTTAAA 1926

Qy    643 Tyr------------------ArgThrGlnSerPheGlyThrHisAlaIleAspLysPhe 656
Db   1927 CATAATGTAGGTGAAGACCCTAATTCAACATTATCTGGTATAGTTTACGTTGACCGAATC 1986

Qy    657 GluPheIleProLeuAsn 662
Db   1987 GAATTCATCCCAGTAGAT 2004
```

RESULT 10
US-11-058-727-81
; Sequence 81, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2022)
US-11-058-727-81

Alignment Scores:

Pred. No.:	7.6e-100	Length:	2022
Score:	925.50	Matches:	243
Percent Similarity:	51.3%	Conservative:	119
Best Local Similarity:	34.4%	Mismatches:	259
Query Match:	25.7%	Indels:	85
DB:	8	Gaps:	26

US-10-782-096-2 (1-682) x US-11-058-727-81 (1-2022)

```
Qy     3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db     4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy    23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db    61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy    43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db   121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy    63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db   163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy    82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db   223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy   102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db   283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy   120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db   343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy   140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db   403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy   160 AlaTrpLeuValAsnLysAsnAspAsp---------AsnArgArgAlaLeu-------- 173
Db   463 GAATGGGAAGAAAATCCAAATGGTTCCCGGTTTCGAAGTCGACAAGCCTTACGAGATGTG 522

Qy   174 ValThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPheLysGlu 193
Db   523 CGAAATCGATTTGAAATCCTGGATAGTTTATTACGCAATATATGCCATCTTTTAGAGTG 582

Qy   194 ArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIle 213
Db   583 ACAAATTTTGAAGTACCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTG 642

Qy   214 LeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArg 233
Db   643 TTATTAAAGGACGCGTCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAAT 702

Qy   234 AspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPhe 253
Db   703 AACTATTATGATCGTCAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGG 762

Qy   254 TyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsn 273
Db   763 TATGAAACTGGTTTAGCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAAC 822

Qy   274 ArgPheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyr 293
Db   823 CAATTCCGTAGAGAAATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTAT 882

Qy   294 AspProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAsp 313
Db   883 GACACACGCACGTACCCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGAT 942

Qy   314 ProValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnPro 333
Db   943 CCACTGGGCGCGGTAAACGTGTCTTCAATTGGTTCC------------TGGTATGAC--- 987

Qy   334 AsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThr 353
Db   988 AAAGCACCTTCTTTCGGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGAT 1047

Qy   354 TrpLeuAsnArgIlePheValTyrThr------ArgThrLeuGlyAsnMetSerAspVal 371
Db  1048 TATATAACGGGACTCACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATA 1107

Qy   372 ArgAsnIleTrpGlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIle 390
Db  1108 AGACAT---TGGGCTGGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTT 1164

Qy   391 ThrHisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsn 410
Db  1165 CAACAAATGTATGGAACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAAT 1224

Qy   411 LeuSerValPheSerIleGluSer---------------LeuAlaArgIleTyrLeuGly 425
Db  1225 TATGATATTTACAAGACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGT 1284

Qy   426 GlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePhe 444
Db  1285 TATACG------------TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTC 1326

Qy   445 AsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeu 464
Db  1327 ATGGTAAACCAATTGAATAATACCAGAAAGACGTTAAAGTAT------------AATCCA 1374

Qy   465 ProSerGlnThrIleLeuSer--------------------GluLeuProGlyLysAsp 477
Db  1375 GTTTCCAAAGATATTATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACT 1434

Qy   478 LysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAsp 497
Db  1435 TCAGATCAACCAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCC 1494

Qy   498 AlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThr 517
Db  1495 GCGACGGGTAACACTACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGA 1545

Qy   518 SerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLys 537
Db  1546 AGTGCAGATTTAAACAATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAA 1605

Qy   538 GlyTrpGlyGlyAsnIleGlyPheValIle------------ProGlyProThrGlyGly 553
Db  1606 TGTTGG---GATAATTTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGG 1662

Qy   554 AsnLeuValLysValSerAspSerTrpHisSerLeuLysVal------------------ 567
Db  1663 GATTTATTACAGTATAATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATAT 1722

Qy   568 ---GlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuVal 586
Db  1723 GGCCTAGCATTAGAAAAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT------ 1776

Qy   587 ThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePhe 606
Db  1777 ---------GATGCAGATATTGTATTGCATGTAAACGATGCTCAGATTCAG--------- 1818

Qy   607 AspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyr 626
Db  1819 ------ATGCCAAAAACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTT 1872

Qy   627 IleAspValProGlyIlePheThrProSerIle------------AsnProLeuIleArg 642
Db  1873 GCAGAT------GCTATCACAACAGTTAATTTAGCAACAGATAGTTCGGTTGCAGTTAAA 1926

Qy   643 Tyr------------------ArgThrGlnSerPheGlyThrHisAlaIleAspLysPhe 656
Db  1927 CATAATTTAGGTGAAGACCCTAATTCAACATTATCTGGTATAGTTTACGTTGACCGAATC 1986

Qy   657 GluPheIleProLeuAsn 662
Db  1987 GAATTCATCCCAGTAGAT 2004
```

RESULT 11

```
US-11-108-389-49
; Sequence 49, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
;  APPLICANT: Andre R. Abad
;  APPLICANT:  Ronald D. Flannagan
;  APPLICANT:  Rafael Herrmann
;  APPLICANT:  Theodore W. Kahn
;  APPLICANT:  Albert L. Lu
;  APPLICANT:  Billy Fred McCutchen
;  APPLICANT:  James K. Presnail
;  APPLICANT:  James F.H. Wong
;  APPLICANT:  Cao-Guo Yu
;  TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
;  TITLE OF INVENTION:  Activity
;  FILE REFERENCE: 35718/291049
;  CURRENT APPLICATION NUMBER: US/11/108,389
;  CURRENT FILING DATE:  2005-04-18
;  PRIOR APPLICATION NUMBER: 60/391,786
;  PRIOR FILING DATE: 2002-06-26
;  PRIOR APPLICATION NUMBER: 60/460,787
;  PRIOR FILING DATE: 2003-04-04
;  PRIOR APPLICATION NUMBER: 10/606,320
;  PRIOR FILING DATE: 2003-06-25
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
;   LENGTH: 2022
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis (mutated)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2022)
US-11-108-389-49

Alignment Scores:
Pred. No.:               7.6e-100      Length:         2022
Score:                   925.50        Matches:        243
Percent Similarity:      51.3%         Conservative:   119
Best Local Similarity:   34.4%         Mismatches:     259
Query Match:             25.7%         Indels:         85
DB:                      8             Gaps:           26

US-10-782-096-2 (1-682) x US-11-108-389-49 (1-2022)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db        4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db       61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      121 GATTATAAAGATTATTTAAAAATGTCT-----------------GCGGGAAATGCTAGT 162

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db      283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db      343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db      403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy      160 AlaTrpLeuValAsnLysAsnAspAsp---------AsnArgArgAlaLeu--------- 173
Db      463 GAATGGGAAGAAAATCCAAATGGTTCCCGGTTTCGAAGTCGACAAGCCTTACGAGATGTG 522

Qy      174 ValThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPheLysGlu 193
Db      523 CGAAATCGATTTGAAATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTG 582

Qy      194 ArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIle 213
Db      583 ACAAATTTTGAAGTACCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTG 642

Qy      214 LeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArg 233
Db      643 TTATTAAAGGACGCGTCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAAT 702

Qy      234 AspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPhe 253
Db      703 AACTATTATGATCGTCAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGG 762

Qy      254 TyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsn 273
Db      763 TATGAAACTGGTTTAGCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAAC 822

Qy      274 ArgPheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyr 293
Db      823 CAATTCCGTAGAGAAATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTAT 882

Qy      294 AspProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAsp 313
Db      883 GACACACGCACGTACCCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGAT 942

Qy      314 ProValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnPro 333
Db      943 CCACTGGGCGCGGTAAACGTGTCTTCAATTGGTTCC------------TGGTATGAC--- 987

Qy      334 AsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThr 353
Db      988 AAAGCACCTTCTTTCGGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGAT 1047

Qy      354 TrpLeuAsnArgIlePheValTyrThr------ArgThrLeuGlyAsnMetSerAspVal 371
Db     1048 TATATAACGGGACTCACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATA 1107

Qy      372 ArgAsnIleTrpGlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIle 390
Db     1108 AGACAT---TGGGCTGGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTT 1164

Qy      391 ThrHisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsn 410
Db     1165 CAACAAATGTATGGAACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAAT 1224

Qy      411 LeuSerValPheSerIleGluSer---------------LeuAlaArgIleTyrLeuGly 425
Db     1225 TATGATATTTACAAGACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGT 1284

Qy      426 GlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePhe 444
Db     1285 TATACG------------TATATATT------TTTGGAATGCCAGAAGTCGAGTTTTTC 1326

Qy      445 AsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeu 464
Db     1327 ATGGTAAACCAATTGAATAATACCAGAAAGACGTTAAAGTAT------------AATCCA 1374

Qy      465 ProSerGlnThrIleLeuSer--------------------GluLeuProGlyLysAsp 477
Db     1375 GTTTCCAAAGATATTATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACT 1434

Qy      478 LysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAsp 497
```

```
Db     1435 TCAGATCAACCAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCC 1494

Qy      498 AlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThr 517
Db     1495 GCGACGGGTAACACTACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGA 1545

Qy      518 SerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLys 537
Db     1546 AGTGCAGATTTAAACAATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAA 1605

Qy      538 GlyTrpGlyGlyAsnIleGlyPheValIle------------ProGlyProThrGlyGly 553
Db     1606 TGTTGG---GATAATTTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGG 1662

Qy      554 AsnLeuValLysValSerAspSerTrpHisSerLeuLysVal------------------ 567
Db     1663 GATTTATTACAGTATAATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATAT 1722

Qy      568 ---GlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuVal 586
Db     1723 GGCCTAGCATTAGAAAAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT------ 1776

Qy      587 ThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePhe 606
Db     1777 ---------GATGCAGATATTGTATTGCATGTAAACGATGCTCAGATTCAG--------- 1818

Qy      607 AspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyr 626
Db     1819 ------ATGCCAAAAACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTT 1872

Qy      627 IleAspValProGlyIlePheThrProSerIle------------AsnProLeuIleArg 642
Db     1873 GCAGAT------GCTATCACAACAGTTAATTTAGCAACAGATAGTTCGGTTGCAGTTAAA 1926

Qy      643 Tyr------------------ArgThrGlnSerPheGlyThrHisAlaIleAspLysPhe 656
Db     1927 CATAATGTAGGTGAAGACCCTAATTCAACATTATCTGGTATAGTTTACGTTGACCGAATC 1986

Qy      657 GluPheIleProLeuAsn 662
Db     1987 GAATTCATCCCAGTAGAT 2004
```

RESULT 12
US-11-108-389-81
; Sequence 81, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2022)
US-11-108-389-81

```
Alignment Scores:
Pred. No.:              7.6e-100     Length:        2022
Score:                  925.50       Matches:       243
Percent Similarity:     51.3%        Conservative:  119
Best Local Similarity:  34.4%        Mismatches:    259
Query Match:            25.7%        Indels:        85
DB:                     8            Gaps:          26

US-10-782-096-2 (1-682) x US-11-108-389-81 (1-2022)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db        4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db       61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db      283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db      343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db      403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy      160 AlaTrpLeuValAsnLysAsnAspAsp---------AsnArgArgAlaLeu--------- 173
Db      463 GAATGGGAAGAAAATCCAAATGGTTCCCGGTTTCGAAGTCGACAAGCCTTACGAGATGTG 522

Qy      174 ValThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPheLysGlu 193
Db      523 CGAAATCGATTTGAAATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTG 582

Qy      194 ArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIle 213
Db      583 ACAAATTTTGAAGTACCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTG 642

Qy      214 LeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArg 233
Db      643 TTATTAAAGGACGCGTCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAAT 702

Qy      234 AspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPhe 253
Db      703 AACTATTATGATCGTCAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGG 762

Qy      254 TyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsn 273
Db      763 TATGAAACTGGTTTAGCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAAC 822

Qy      274 ArgPheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyr 293
```

```
                :::||||||  :::|||||||||  |||||||||:::         ||||||||||||||
Db        823 CAATTCCGTAGAGAAATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTAT 882

Qy        294 AspProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAsp 313
              |||   |||   ||||||:::     |||   :::||||||||||||||||||||||||
Db        883 GACACACGCACGTACCCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGAT 942

Qy        314 ProValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnPro 333
              |||:::|||         |||         |||         ||||||:::
Db        943 CCACTGGGCGCGGTAAACGTGTCTTCAATTGGTTCC------------TGGTATGAC--- 987

Qy        334 AsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThr 353
                       :::|||    :::|||::::::      |||   |||
Db        988 AAAGCACCTTCTTTCGGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGAT 1047

Qy        354 TrpLeuAsnArgIlePheValTyrThr------ArgThrLeuGlyAsnMetSerAspVal 371
              ::::::   :::   |||||||||      |||::::::   :::         :::
Db       1048 TATATAACGGGACTCACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATA 1107

Qy        372 ArgAsnIleTrpGlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIle 390
              |||:::   |||   ||||||               :::   ||||||    :::
Db       1108 AGACAT---TGGGCTGGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTT 1164

Qy        391 ThrHisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsn 410
                       :::|||      ::::::            |||:::|||   |||
Db       1165 CAACAAATGTATGGAACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAAT 1224

Qy        411 LeuSerValPheSerIleGluSer---------------LeuAlaArgIleTyrLeuGly 425
                 ::::::         |||               |||      :::|||   |||
Db       1225 TATGATATTTACAAGACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGT 1284

Qy        426 GlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePhe 444
                 |||            ||||||      :::|||:::      |||      |||
Db       1285 TATACG------------TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTC 1326

Qy        445 AsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeu 464
                 :::   :::||||||         :::|||:::|||               |||
Db       1327 ATGGTAAACCAATTGAATAATACCAGAAAGACGTTAAAGTAT------------AATCCA 1374

Qy        465 ProSerGlnThrIleLeuSer--------------------GluLeuProGlyLysAsp 477
                 |||:::   |||::::::                    |||||||||   :::
Db       1375 GTTTCCAAAGATATTATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACT 1434

Qy        478 LysProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAsp 497
                 :::||||||         :::||||||||||||   :::|||::::::
Db       1435 TCAGATCAACCAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCC 1494

Qy        498 AlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThr 517
              |||      ::::::::|||   :::                  |||   |||   |||
Db       1495 GCGACGGGTAACACTACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGA 1545

Qy        518 SerMetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLys 537
              |||   |||   ||||||   :::         ||||||||||||||||||   |||||||||
Db       1546 AGTGCAGATTTAAACAATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAA 1605

Qy        538 GlyTrpGlyGlyAsnIleGlyPheValIle------------ProGlyProThrGlyGly 553
                 |||   |||:::   ||||||                  ||||||   |||||||||
Db       1606 TGTTGG---GATAATTTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGG 1662

Qy        554 AsnLeuValLysValSerAspSerTrpHisSerLeuLysVal------------------ 567
              :::|||::::::   :::   |||   |||:::
Db       1663 GATTTATTACAGTATAATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATAT 1722

Qy        568 ---GlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuVal 586
                    |||   ::::::         ||||||:::|||||||||||||||
Db       1723 GGCCTAGCATTAGAAAAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT------ 1776

Qy        587 ThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePhe 606
                    ||||||   |||   |||         :::   |||
Db       1777 ---------GATGCAGATATTGTATTGCATGTAAACGATGCTCAGATTCAG--------- 1818

Qy        607 AspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyr 626
                       :::      |||      |||         |||:::
Db       1819 ------ATGCCAAAAACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTT 1872

Qy        627 IleAspValProGlyIlePheThrProSerIle------------AsnProLeuIleArg 642
                 |||         |||   |||   ::::::            :::      ::::::
Db       1873 GCAGAT------GCTATCACAACAGTTAATTTAGCAACAGATAGTTCGGTTGCAGTTAAA 1926

Qy        643 Tyr-----------------ArgThrGlnSerPheGlyThrHisAlaIleAspLysPhe 656
              :::                 :::            |||         :::|||:::
Db       1927 CATAATTTAGGTGAAGACCCTAATTCAACATTATCTGGTATAGTTTACGTTGACCGAATC 1986

Qy        657 GluPheIleProLeuAsn 662
              ||||||||||||::::::
Db       1987 GAATTCATCCCAGTAGAT 2004
```

RESULT 13
US-11-058-727-7
; Sequence 7, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 2022
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis (mutated)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2022)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: NGSR.N1218-1
US-11-058-727-7

Alignment Scores:

Pred. No.:	8.72e-100	Length:	2022
Score:	925.00	Matches:	239
Percent Similarity:	50.5%	Conservative:	116
Best Local Similarity:	34.0%	Mismatches:	268
Query Match:	25.7%	Indels:	80
DB:	8	Gaps:	23

US-10-782-096-2 (1-682) x US-11-058-727-7 (1-2022)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                 |||      |||:::||||||||||||:::::::||||||      ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||      |||||||||   |||   :::|||         ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
```

```
Db    121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy     63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81

Db    163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy     82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101

Db    223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy    102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119

Db    283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy    120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139

Db    343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy    140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159

Db    403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy    160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---------------ArgAlaLeuVal 174

Db    463 GAATGGGAAGAAAATCCAAATGGTTCAAGAAATGGTTCCCGGGCCTTACGAGATGTGCGA 522

Qy    175 ThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArg 194

Db    523 AATCGATTTGAAATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACA 582

Qy    195 AsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeu 214

Db    583 AATTTTGAAGTACCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTA 642

Qy    215 LeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAsp 234

Db    643 TTAAAGGACGCGTCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAAC 702

Qy    235 AsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyr 254

Db    703 TATTATGATCGTCAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTAT 762

Qy    255 AsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArg 274

Db    763 GAAACTGGTTTAGCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAA 822

Qy    275 PheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAsp 294

Db    823 TTCCGTAGAGAAATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGAC 882

Qy    295 ProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspPro 314

Db    883 ACACGCACGTACCCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCA 942

Qy    315 ValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsn 334

Db    943 CTGGGCGCGGTAAACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAA 987

Qy    335 AsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrp 354

Db    988 GCACCTTCTTTCGGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTAT 1047

Qy    355 LeuAsnArgIlePheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArg 372

Db   1048 ATAACGGGACTCACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGA 1107

Qy    373 AsnIleTrpGlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThr 391

Db   1108 CAT---TGGGCTGGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAA 1164

Qy    392 HisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeu 411

Db   1165 CAAATGTATGGAACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTAT 1224

Qy    412 SerValPheSerIleGluSer---------------LeuAlaArgIleTyrLeuGlyGly 426

Db   1225 GATATTTACAAGACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTAT 1284

Qy    427 ThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsn 445

Db   1285 ACG------------TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATG 1326

Qy    446 ThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuPro 465

Db   1327 GTAAACCAATTGAATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTT 1374

Qy    466 SerGlnThrIleLeuSer--------------------GluLeuProGlyLysAspLys 478

Db   1375 TCCAAAGATATTATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCA 1434

Qy    479 ProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAla 498

Db   1435 GATCAACCAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCG 1494

Qy    499 ArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSer 518

Db   1495 ACGGGTAACACTACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGT 1545

Qy    519 MetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGly 538

Db   1546 GCAGATTTAAACAATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGT 1605

Qy    539 TrpGlyGlyAsnIleGlyPheValIle------------ProGlyProThrGlyGlyAsn 554

Db   1606 TGG---GATAATTTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGAT 1662

Qy    555 LeuValLysValSerAspSerTrpHisSerLeuLysVal--------------------- 567

Db   1663 TTATTACAGTATAATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGC 1722

Qy    568 GlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThr 587

Db   1723 CTAGCATTAGAAAAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT--------- 1773

Qy    588 HisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAsp 607

Db   1774 ------GATGCAGATATTGTATTGCATGTAAACGATGCTCAGATTCAG------------ 1815

Qy    608 CysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIle 627

Db   1816 ---ATGCCAAAAACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCA 1872

Qy    628 AspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArg--------- 644

Db   1873 GATGCTATCACAACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTA 1932

Qy    645 ---------------ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIle 659

Db   1933 GGTGAAGACCCTAATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATC 1992

Qy    660 ProLeuAsn 662

Db   1993 CCAGTAGAT 2001
```

RESULT 14
US-11-058-727-25
; Sequence 25, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen

```
;  APPLICANT:  James K. Presnail
;  APPLICANT:  James F.H. Wong
;  APPLICANT:  Cao-Guo Yu
;  TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
;  TITLE OF INVENTION:  Activity
;  FILE REFERENCE: 35718/287809
;  CURRENT APPLICATION NUMBER: US/11/058,727
;  CURRENT FILING DATE:  2005-02-15
;  PRIOR APPLICATION NUMBER: 60/391,786
;  PRIOR FILING DATE: 2002-06-26
;  PRIOR APPLICATION NUMBER: 60/460,787
;  PRIOR FILING DATE: 2003-04-04
;  PRIOR APPLICATION NUMBER: 10/606,320
;  PRIOR FILING DATE: 2003-06-25
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
;   LENGTH: 2022
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis (mutated)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2022)
US-11-058-727-25

Alignment Scores:
Pred. No.:              8.72e-100      Length:         2022
Score:                  925.00         Matches:        239
Percent Similarity:     50.5%          Conservative:   116
Best Local Similarity:  34.0%          Mismatches:     268
Query Match:            25.7%          Indels:         80
DB:                     8              Gaps:           23

US-10-782-096-2 (1-682) x US-11-058-727-25 (1-2022)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---------------ArgAlaLeuVal 174
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAAATGGTTCCCGGGCCTTACGAGATGTGCGA 522

Qy        175 ThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArg 194
Db        523 AATCGATTTGAAATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACA 582

Qy        195 AsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeu 214
Db        583 AATTTTGAAGTACCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTA 642

Qy        215 LeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAsp 234
Db        643 TTAAAGGACGCGTCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAAC 702

Qy        235 AsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyr 254
Db        703 TATTATGATCGTCAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTAT 762

Qy        255 AsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArg 274
Db        763 GAAACTGGTTTAGCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAA 822

Qy        275 PheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAsp 294
Db        823 TTCCGTAGAGAAATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGAC 882

Qy        295 ProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspPro 314
Db        883 ACACGCACGTACCCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCA 942

Qy        315 ValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsn 334
Db        943 CTGGGCGCGGTAAACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAA 987

Qy        335 AsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrp 354
Db        988 GCACCTTCTTTCGGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTAT 1047

Qy        355 LeuAsnArgIlePheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArg 372
Db       1048 ATAACGGGACTCACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGA 1107

Qy        373 AsnIleTrpGlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThr 391
Db       1108 CAT---TGGCTGGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAA 1164

Qy        392 HisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeu 411
Db       1165 CAAATGTATGGAACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTAT 1224

Qy        412 SerValPheSerIleGluSer---------------LeuAlaArgIleTyrLeuGlyGly 426
Db       1225 GATATTTACAAGACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTAT 1284

Qy        427 ThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsn 445
Db       1285 ACG------------TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATG 1326

Qy        446 ThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuPro 465
Db       1327 GTAAACCAATTGAATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTT 1374

Qy        466 SerGlnThrIleLeuSer--------------------GluLeuProGlyLysAspLys 478
Db       1375 TCCAAAGATATTATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCA 1434

Qy        479 ProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAla 498
Db       1435 GATCAACCAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCG 1494

Qy        499 ArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSer 518
Db       1495 ACGGGTAACACTACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGT 1545

Qy        519 MetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGly 538
Db       1546 GCAGATTTAAACAATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGT 1605
```

```
Qy        539 TrpGlyGlyAsnIleGlyPheValIle------------ProGlyProThrGlyGlyAsn 554
Db       1606 TGG---GATAATTTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGAT 1662

Qy        555 LeuValLysValSerAspSerTrpHisSerLeuLysVal--------------------- 567
Db       1663 TTATTACAGTATAATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGC 1722

Qy        568 GlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThr 587
Db       1723 CTAGCATTAGAAAAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT--------- 1773

Qy        588 HisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAsp 607
Db       1774 ------GATGCAGATATTGTATTGCATGTAAACGATGCTCAGATTCAG------------ 1815

Qy        608 CysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIle 627
Db       1816 ---ATGCCAAAAACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCA 1872

Qy        628 AspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArg--------- 644
Db       1873 GATGCTATCACAACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATGTA 1932

Qy        645 ---------------ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIle 659
Db       1933 GGTGAAGACCCTAATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATC 1992

Qy        660 ProLeuAsn 662
Db       1993 CCAGTAGAT 2001
```

RESULT 15
US-11-058-727-29
; Sequence 29, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (mutated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2022)
US-11-058-727-29

Alignment Scores:

Pred. No.:	8.72e-100	Length:	2022
Score:	925.00	Matches:	239
Percent Similarity:	50.5%	Conservative:	116
Best Local Similarity:	34.0%	Mismatches:	268
Query Match:	25.7%	Indels:	80
DB:	8	Gaps:	23

US-10-782-096-2 (1-682) x US-11-058-727-29 (1-2022)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---------------ArgAlaLeuVal 174
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAAATGGTTCCCGGGCCTTACGAGATGTGCGA 522

Qy        175 ThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArg 194
Db        523 AATCGATTTGAAATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACA 582

Qy        195 AsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeu 214
Db        583 AATTTTGAAGTACCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTA 642

Qy        215 LeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAsp 234
Db        643 TTAAAGGACGCGTCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAAC 702

Qy        235 AsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyr 254
Db        703 TATTATGATCGTCAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTAT 762

Qy        255 AsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArg 274
Db        763 GAAACTGGTTTAGCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAA 822

Qy        275 PheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAsp 294
Db        823 TTCCGTAGAGAAATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGAC 882

Qy        295 ProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspPro 314
Db        883 ACACGCACGTACCCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCA 942

Qy        315 ValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsn 334
Db        943 CTGGGCGCGGTAAACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAA 987
```

```
Qy        335 AsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrp 354
                    :::|||      :::|||::::::      |||   |||            :::
Db        988 GCACCTTCTTTCGGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTAT 1047

Qy        355 LeuAsnArgIlePheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArg 372
              :::      :::   ||||||||      |||::::::   :::         :::|||
Db       1048 ATAACGGGACTCACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGA 1107

Qy        373 AsnIleTrpGlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThr 391
              :::   |||   ||||||   :::            :::   ||||||   :::
Db       1108 CAT---TGGGCTGGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAA 1164

Qy        392 HisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeu 411
                    :::|||         ::::::            |||:::|||   |||
Db       1165 CAAATGTATGGAACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTAT 1224

Qy        412 SerValPheSerIleGluSer---------------LeuAlaArgIleTyrLeuGlyGly 426
                 ::::::      |||               |||   :::|||   |||
Db       1225 GATATTTACAAGACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTAT 1284

Qy        427 ThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsn 445
              |||            ||||||         :::|||:::      |||      |||
Db       1285 ACG------------TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATG 1326

Qy        446 ThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuPro 465
                 :::   :::||||||         :::|||:::|||            |||
Db       1327 GTAAACCAATTGAATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTT 1374

Qy        466 SerGlnThrIleLeuSer--------------------GluLeuProGlyLysAspLys 478
              |||:::   |||::::::                    |||||||||   :::
Db       1375 TCCAAAGATATTATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCA 1434

Qy        479 ProArgProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAla 498
                 :::||||||         :::|||||||||||||   :::|||::::::      |||
Db       1435 GATCAACCAAATTATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCG 1494

Qy        499 ArgArgSerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSer 518
                    ::::::::|||   :::                |||   |||   |||   |||
Db       1495 ACGGGTAACACTACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGT 1545

Qy        519 MetAspArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGly 538
                 |||   ||||||   :::      ||||||||||||||||   |||||||||
Db       1546 GCAGATTTAAACAATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGT 1605

Qy        539 TrpGlyGlyAsnIleGlyPheValIle------------ProGlyProThrGlyGlyAsn 554
              |||   |||:::   ||||||            ||||||   |||||||||:::
Db       1606 TGG---GATAATTTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGAT 1662

Qy        555 LeuValLysValSerAspSerTrpHisSerLeuLysVal--------------------- 567
              |||::::::   :::   |||   |||:::
Db       1663 TTATTACAGTATAATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGC 1722

Qy        568 GlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThr 587
                 |||   ::::::         ||||||:::||||||||||||||
Db       1723 CTAGCATTAGAAAAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT--------- 1773

Qy        588 HisGlyAspAlaIlePheValGluHisSerGlySerSerHisIleValSerPhePheAsp 607
                    ||||||      |||   |||         :::   |||
Db       1774 ------GATGCAGATATTGTATTGCATGTAAACGATGCTCAGATTCAG------------ 1815

Qy        608 CysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIle 627
                       :::      |||         |||               |||:::
Db       1816 ---ATGCCAAAAACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCA 1872

Qy        628 AspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArg--------- 644
              |||                        :::   :::      :::      :::
Db       1873 GATGCTATCACAACATTAAATTTAGCAACAGATAGTTCGCTAGCAGTGAAACATAATGTA 1932

Qy        645 ---------------ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIle 659
                             :::         |||         :::|||:::   |||||||||
Db       1933 GGTGAAGACCCTAATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATC 1992

Qy        660 ProLeuAsn 662
              |||::::::
Db       1993 CCAGTAGAT 2001
```

Search completed: January 21, 2006, 04:28:55
Job time : 366 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          January 20, 2006, 22:28:18 ; Search time 9349 Seconds
                                              (without alignments)
                                              4146.672 Million cell updates/sec

Title:           US-10-782-096-2
Perfect score:   3597
Sequence:        1 MNSYKNKNEYEMLDALRINS..........TFPNQSLEKREQEVNDLFIN 682

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters:      11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10782096/runat_20012006_095214_24394/app_query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10782096_@CGN_1_1_5961_@runat_20012006_095214_24394 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_in:*
                 3:  gb_env:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pr:*
                 9:  gb_ro:*
                 10:  gb_sts:*
                 11:  gb_sy:*
                 12:  gb_un:*
                 13:  gb_vi:*
                 14:  gb_htg:*
                 15:  gb_pl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3597	100.0	2049	6	CQ868318	CQ868318 Sequence
2	3536	98.3	2016	6	CQ868320	CQ868320 Sequence
3	3488	97.0	1986	6	CQ868322	CQ868322 Sequence
4	948	26.4	3621	6	CS130946	CS130946 Sequence
5	948	26.4	3621	6	CS132867	CS132867 Sequence
6	948	26.4	3621	6	AX543924	AX543924 Sequence
7	948	26.4	4874	6	CS130962	CS130962 Sequence
8	948	26.4	4874	6	AX543950	AX543950 Sequence
9	942.5	26.2	3507	1	BTU04365	U04365 Bacillus th
10	942.5	26.2	3507	6	I25972	I25972 Sequence 3
11	940.5	26.1	3504	6	BD133574	BD133574 Protein h
12	940.5	26.1	3690	6	BD133575	BD133575 Protein h
13	930.5	25.9	3435	1	AB089299	AB089299 Bacillus
14	926	25.7	2003	6	AX543928	AX543928 Sequence
15	926	25.7	2010	6	CS130950	CS130950 Sequence
16	926	25.7	2010	6	CS130956	CS130956 Sequence
17	926	25.7	2010	6	AX543932	AX543932 Sequence
18	926	25.7	2010	6	AX543938	AX543938 Sequence
19	925.5	25.7	2022	6	CS130994	CS130994 Sequence
20	925.5	25.7	2022	6	CS131026	CS131026 Sequence
21	925	25.7	2022	6	CS130952	CS130952 Sequence
22	925	25.7	2022	6	CS130970	CS130970 Sequence
23	925	25.7	2022	6	CS130974	CS130974 Sequence
24	925	25.7	2022	6	CS130978	CS130978 Sequence
25	925	25.7	2022	6	CS131014	CS131014 Sequence
26	925	25.7	2022	6	AX543934	AX543934 Sequence
27	924.5	25.7	2022	6	CS130988	CS130988 Sequence
28	924.5	25.7	2022	6	CS131020	CS131020 Sequence
29	924	25.7	2025	6	CS130990	CS130990 Sequence
30	924	25.7	2025	6	CS131022	CS131022 Sequence
31	924	25.7	3633	6	CS130948	CS130948 Sequence
32	924	25.7	3633	6	AX543926	AX543926 Sequence
33	924	25.7	6613	6	CS130963	CS130963 Sequence
34	924	25.7	6613	6	AX543951	AX543951 Sequence
35	923	25.7	2022	6	AX543944	AX543944 Sequence
36	920.5	25.6	2013	6	AX543946	AX543946 Sequence
37	920	25.6	2022	6	AX543962	AX543962 Sequence
38	920	25.6	2025	6	CS130992	CS130992 Sequence
39	920	25.6	2025	6	CS131024	CS131024 Sequence
40	919.5	25.6	2013	6	AX543966	AX543966 Sequence
41	916	25.5	2022	6	CS130966	CS130966 Sequence
42	916	25.5	2022	6	CS131008	CS131008 Sequence
43	916	25.5	2022	6	CS131010	CS131010 Sequence
44	916	25.5	2022	6	CS131012	CS131012 Sequence
45	916	25.5	2025	6	CS130986	CS130986 Sequence

ALIGNMENTS

```
RESULT 1
CQ868318
LOCUS       CQ868318                2049 bp    DNA     linear   PAT 13-SEP-2004
DEFINITION  Sequence 19 from Patent WO2004074462.
ACCESSION   CQ868318
VERSION     CQ868318.1  GI:51998364
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Carozzi,N., Hargiss,T., Koziel,M.G., Duck,N.B. and Carr,B.
  TITLE     Delta-endotoxin genes and methods for their use
  JOURNAL   Patent: WO 2004074462-A 19 02-SEP-2004;
            Athenix Corporation (US)
FEATURES             Location/Qualifiers
     source          1..2049
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
     CDS             1..2049
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAH33954.1"
                     /db_xref="GI:51998365"
```

```
                     /translation="MNSYKNKNEYEMLDALRINSNMSNCYPRYPLAKDPQMTMRNTNY
                     KEWLNMCDSNTQFIGDISTYSSPEAALSVRDAVLTGINSVGTILSNLGVPLASQSFGI
                     ISRLIGILWAGPDPFEALMVLVEELIKKSIDQRVRENALRELEGLQGIMRLYQTRLQA
                     WLVNKNDDNRRALVTQYAIVDNFFEKNMPKFKERNFEILLLPVYAQAANLHLILLRDA
                     DYFGAQWQLGDDEIRDNYIRLQGLIREYKDHCITFYNQGLNQFNRSNAQDWVSFNRFR
                     TDMTLTVLDLAILFPNYDPRRYPLAVKTELTREVYTDPVGFTGVLESGGRTYPWYNPN
                     NTTFTAMENNARRRPSYTTWLNRIFVYTRTLGNMSDVRNIWGGHTLVENGNDGSEITH
                     NFGKTDSITPIQYFNFANLSVFSIESLARIYLGGTEANNYITSQYGVSRVIFNTSNIN
                     NVPGSLRYEVPANLPSQTILSELPGKDKPRPNAGDFSHRLSYISNFDARRSSSGGIVS
                     LLTFGWAHTSMDRNNRLEPDKITQIDAVKGWGGNIGFVIPGPTGGNLVKVSDSWHSLK
                     VQAPQRQTSYRIRLRYACLVTHGDAIFVEHSGSSHIVSFFDCSNSSGRPSNTLLESDF
                     RYIDVPGIFTPSINPLIRYRTQSFGTHAIDKFEFIPLNTFPNQSLEKREQEVNDLFIN
                     "
ORIGIN

Alignment Scores:
Pred. No.:              1.9e-259        Length:         2049
Score:                  3597.00         Matches:        682
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-782-096-2 (1-682) x CQ868318 (1-2049)

Qy          1 MetAsnSerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATTCATATAAAAATAAAAATGAATATGAAATGTTGGATGCTTTACGAATCAACTCT 60

Qy         21 AsnMetSerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATATGTCTAATTGTTATCCAAGGTATCCACTAGCAAAAGATCCACAAATGACTATGCGA 120

Qy         41 AsnThrAsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAsp 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACACGAACTATAAAGAATGGCTAAATATGTGTGATTCAAATACACAATTTATTGGTGAT 180

Qy         61 IleSerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGly 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAAGCACGTATTCTAGCCCTGAAGCTGCTTTAAGTGTACGAGATGCTGTTTTAACGGGT 240

Qy         81 IleAsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTAACAGTGTAGGGACTATACTTTCGAATTTAGGGGTCCCTTTGGCAAGTCAATCATTT 300

Qy        101 GlyIleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAspProPheGluAlaLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAATAATTAGTAGGCTAATAGGTATTTTATGGGCAGGGCCTGATCCATTTGAAGCACTT 360

Qy        121 MetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGTTCTTGTTGAAGAGCTTATTAAGAAAAGTATAGATCAGCGTGTAAGAGAAAATGCT 420

Qy        141 LeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAla 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTAGAGAGCTAGAAGGTTTACAGGGAATTATGAGACTATATCAAACTAGACTGCAAGCA 480

Qy        161 TrpLeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThrGlnTyrAlaIleVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGCTAGTTAACAAGAATGATGACAATCGGAGGGCACTAGTAACGCAGTATGCAATTGTT 540

Qy        181 AspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATAACTTTTTCGAAAAGAATATGCCAAAATTCAAGGAAAGAAACTTTGAAATTTTATTG 600

Qy        201 LeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTACCAGTATATGCACAAGCCGCGAATTTGCATTTAATTTTATTAAGAGATGCTGATTAT 660

Qy        221 PheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTGGAGCACAGTGGCAATTAGGTGATGATGAAATTCGTGATAATTATATCAGACTACAA 720

Qy        241 GlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGln 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGACTGATTAGAGAATATAAAGATCATTGTATAACATTCTATAACCAGGGTTTAAATCAA 780

Qy        261 PheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTAATCGCTCAAATGCTCAAGATTGGGTGAGCTTTAATAGGTTTCGTACAGATATGACA 840

Qy        281 LeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTAACAGTATTAGATCTCGCAATATTATTTCCAAACTATGATCCACGTAGGTATCCATTA 900

Qy        301 AlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCAGTAAAAACGGAATTGACTAGGGAAGTTTATACAGATCCAGTAGGGTTTACTGGGGTA 960

Qy        321 LeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAla 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAGAAAGTGGAGGTAGGACTTACCCTTGGTATAATCCTAATAATACAACCTTTACTGCT 1020

Qy        341 MetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheVal 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATGGAAAATAACGCAAGACGACGTCCTTCTTATACCACTTGGCTTAATCGTATTTTTGTA 1080

Qy        361 TyrThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHisThrLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TATACAAGGACTCTAGGTAATATGTCTGATGTGAGAAATATTTGGGGAGGGCATACATTA 1140

Qy        381 ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAspSerIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTTGAAAATGGAAATGATGGTTCTGAAATAACCCATAACTTTGGTAAAACTGATTCTATT 1200

Qy        401 ThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAla 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCCTATTCAATATTTTAATTTCGCGAACCTTTCTGTTTTCAGTATTGAGTCACTTGCT 1260

Qy        421 ArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CGTATATATTTAGGAGGAACAGAGGCTAATAATTATATTACTAGTCAGTATGGAGTCTCG 1320

Qy        441 ArgValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluVal 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAGTTATTTTTAATACATCAAATATAAATAATGTACCTGGATCTTTAAGATACGAAGTG 1380

Qy        461 ProAlaAsnLeuProSerGlnThrIleLeuSerGluLeuProGlyLysAspLysProArg 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTGCTAATCTTCCATCCCAAACTATATTATCAGAATTACCAGGAAAGGATAAGCCAAGA 1440

Qy        481 ProAsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArg 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCAAACGCAGGAGATTTCAGCCATAGATTATCTTATATATCAAATTTTGATGCACGGCGA 1500

Qy        501 SerSerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAsp 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGTAGTTCAGGCGGTATTGTTAGTCTTTTAACGTTTGGTTGGGCACATACCAGTATGGAT 1560

Qy        521 ArgAsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CGTAATAATCGTCTTGAACCAGATAAAATTACTCAAATAGATGCAGTTAAAGGTTGGGGG 1620

Qy        541 GlyAsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAsp 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGAATATCGGGTTTGTCATCCCAGGACCTACTGGGGGGAATTTGGTAAAAGTCAGTGAT 1680

Qy        561 SerTrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeu 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGTTGGCATTCACTTAAAGTTCAAGCACCACAAAGACAAACAAGTTATCGTATTCGTTTG 1740

Qy        581 ArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSer 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGTTATGCTTGTTTAGTTACCCATGGGGATGCTATTTTTGTAGAACACAGCGGCAGTAGT 1800

Qy        601 HisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeu 620
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CATATAGTTTCATTTTTGATTGCTCAAATTCATCAGGTCGTCCATCAAACACTCTTCTA 1860

Qy        621 GluSerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GAGAGTGATTTTCGCTATATTGATGTTCCAGGTATTTTTACACCATCAATAAATCCCTTA 1920

Qy        641 IleArgTyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIlePro 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ATAAGATATAGAACACAAAGCTTTGGTACCCACGCGATAGACAAATTTGAATTTATTCCA 1980

Qy        661 LeuAsnThrPheProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPhe 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CTTAACACTTTTCCGAATCAATCATTAGAAAAAAGAGAACAGGAAGTAAATGATCTATTT 2040

Qy        681 IleAsn 682
              ||||||
Db       2041 ATCAAT 2046
```

RESULT 2
CQ868320

```
LOCUS       CQ868320                2016 bp    DNA     linear   PAT 13-SEP-2004
DEFINITION  Sequence 21 from Patent WO2004074462.
ACCESSION   CQ868320
VERSION     CQ868320.1  GI:51998366
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Carozzi,N., Hargiss,T., Koziel,M.G., Duck,N.B. and Carr,B.
  TITLE     Delta-endotoxin genes and methods for their use
  JOURNAL   Patent: WO 2004074462-A 21 02-SEP-2004;
            Athenix Corporation (US)
FEATURES             Location/Qualifiers
     source          1..2016
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
     CDS             1..2016
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAH33955.1"
                     /db_xref="GI:51998367"
                     /translation="MLDALRINSNMSNCYPRYPLAKDPQMTMRNTNYKEWLNMCDSNT
                     QFIGDISTYSSPEAALSVRDAVLTGINSVGTILSNLGVPLASQSFGIISRLIGILWAG
                     PDPFEALMVLVEELIKKSIDQRVRENALRELEGLQGIMRLYQTRLQAWLVNKNDDNRR
                     ALVTQYAIVDNFFEKNMPKFKERNFEILLPVYAQAANLHLILLRDADYFGAQWQLGD
                     DEIRDNYIRLQGLIREYKDHCITFYNQGLNQFNRSNAQDWVSFNRFRTDMTLTVLDLA
                     ILFPNYDPRRYPLAVKTELTREVYTDPVGFTGVLESGGRTYPWYNPNNTFTAMENNA
                     RRRPSYTTWLNRIFVYTRTLGNMSDVRNIWGGHTLVENGNDGSEITHNFGKTDSITPI
                     QYFNFANLSVFSIESLARIYLGGTEANNYITSQYGVSRVIFNTSNINNVPGSLRYEVP
                     ANLPSQTILSELPGKDKPRPNAGDFSHRLSYISNFDARRSSSGGIVSLLTFGWAHTSM
                     DRNNRLEPDKITQIDAVKGWGGNIGFVIPGPTGGNLVKVSDSWHSLKVQAPQRQTSYR
                     IRLRYACLVTHGDAIFVEHSGSSHIVSFFDCSNSSGRPSNTLLESDFRYIDVPGIFTP
                     SINPLIRYRTQSFGTHAIDKFEFIPLNTFPNQSLEKREQEVNDLFIN"
ORIGIN
```

Alignment Scores:

```
Pred. No.:              6.82e-255     Length:         2016
Score:                  3536.00       Matches:        671
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            98.3%         Indels:         0
DB:                     6             Gaps:           0
```

US-10-782-096-2 (1-682) x CQ868320 (1-2016)

```
Qy         12 MetLeuAspAlaLeuArgIleAsnSerAsnMetSerAsnCysTyrProArgTyrProLeu 31
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTTGGATGCTTTACGAATCAACTCTAATATGTCTAATTGTTATCCAAGGTATCCACTA 60

Qy         32 AlaLysAspProGlnMetThrMetArgAsnThrAsnTyrLysGluTrpLeuAsnMetCys 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAAAAGATCCACAAATGACTATGCGAAACACGAACTATAAAGAATGGCTAAATATGTGT 120

Qy         52 AspSerAsnThrGlnPheIleGlyAspIleSerThrTyrSerSerProGluAlaAlaLeu 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATTCAAATACACAATTTATTGGTGATATAAGCACGTATTCTAGCCCTGAAGCTGCTTTA 180

Qy         72 SerValArgAspAlaValLeuThrGlyIleAsnSerValGlyThrIleLeuSerAsnLeu 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTGTACGAGATGCTGTTTTAACGGGTATTAACAGTGTAGGGACTATACTTTCGAATTTA 240

Qy         92 GlyValProLeuAlaSerGlnSerPheGlyIleIleSerArgLeuIleGlyIleLeuTrp 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGGTCCCTTTGGCAAGTCAATCATTTGGAATAATTAGTAGGCTAATAGGTATTTTATGG 300

Qy        112 AlaGlyProAspProPheGluAlaLeuMetValLeuValGluGluLeuIleLysLysSer 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCAGGGCCTGATCCATTTGAAGCACTTATGGTTCTTGTTGAAGAGCTTATTAAGAAAAGT 360

Qy        132 IleAspGlnArgValArgGluAsnAlaLeuArgGluLeuGluGlyLeuGlnGlyIleMet 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATAGATCAGCGTGTAAGAGAAAATGCTCTTAGAGAGCTAGAAGGTTTACAGGGAATTATG 420

Qy        152 ArgLeuTyrGlnThrArgLeuGlnAlaTrpLeuValAsnLysAsnAspAspAsnArgArg 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGACTATATCAAACTAGACTGCAAGCATGGCTAGTTAACAAGAATGATGACAATCGGAGG 480

Qy        172 AlaLeuValThrGlnTyrAlaIleValAspAsnPhePheGluLysAsnMetProLysPhe 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACTAGTAACGCAGTATGCAATTGTTGATAACTTTTTCGAAAAGAATATGCCAAAATTC 540

Qy        192 LysGluArgAsnPheGluIleLeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHis 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGGAAAGAAACTTTGAAATTTTATTGTTACCAGTATATGCACAAGCCGCGAATTTGCAT 600

Qy        212 LeuIleLeuLeuArgAspAlaAspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGlu 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTAATTTTATTAAGAGATGCTGATTATTTTGGAGCACAGTGGCAATTAGGTGATGATGAA 660

Qy        232 IleArgAspAsnTyrIleArgLeuGlnGlyLeuIleArgGluTyrLysAspHisCysIle 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTCGTGATAATTATATCAGACTACAAGGACTGATTAGAGAATATAAAGATCATTGTATA 720

Qy        252 ThrPheTyrAsnGlnGlyLeuAsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSer 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACATTCTATAACCAGGGTTTAAATCAATTTAATCGCTCAAATGCTCAAGATTGGGTGAGC 780

Qy        272 PheAsnArgPheArgThrAspMetThrLeuThrValLeuAspLeuAlaIleLeuPhePro 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTAATAGGTTTCGTACAGATATGACATTAACAGTATTAGATCTCGCAATATTATTTCCA 840

Qy        292 AsnTyrAspProArgArgTyrProLeuAlaValLysThrGluLeuThrArgGluValTyr 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACTATGATCCACGTAGGTATCCATTAGCAGTAAAAACGGAATTGACTAGGGAAGTTTAT 900

Qy        312 ThrAspProValGlyPheThrGlyValLeuGluSerGlyGlyArgThrTyrProTrpTyr 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACAGATCCAGTAGGGTTTACTGGGGTATTAGAAAGTGGAGGTAGGACTTACCCTTGGTAT 960

Qy        332 AsnProAsnAsnThrThrPheThrAlaMetGluAsnAsnAlaArgArgArgProSerTyr 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATCCTAATAATACAACCTTTACTGCTATGGAAAATAACGCAAGACGACGTCCTTCTTAT 1020

Qy        352 ThrThrTrpLeuAsnArgIlePheValTyrThrArgThrLeuGlyAsnMetSerAspVal 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACCACTTGGCTTAATCGTATTTTTGTATATACAAGGACTCTAGGTAATATGTCTGATGTG 1080

Qy        372 ArgAsnIleTrpGlyGlyHisThrLeuValGluAsnGlyAsnAspGlySerGluIleThr 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGAAATATTTGGGGAGGGCATACATTAGTTGAAAATGGAAATGATGGTTCTGAAATAACC 1140
```

```
Qy        392 HisAsnPheGlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeu 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CATAACTTTGGTAAAACTGATTCTATTACTCCTATTCAATATTTTAATTTCGCGAACCTT 1200

Qy        412 SerValPheSerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsn 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTGTTTTCAGTATTGAGTCACTTGCTCGTATATATTTAGGAGGAACAGAGGCTAATAAT 1260

Qy        432 TyrIleThrSerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsn 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TATATTACTAGTCAGTATGGAGTCTCGAGAGTTATTTTTAATACATCAAATATAAATAAT 1320

Qy        452 ValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeuSer 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTACCTGGATCTTTAAGATACGAAGTGCCTGCTAATCTTCCATCCCAAACTATATTATCA 1380

Qy        472 GluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPheSerHisArgLeuSer 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GAATTACCAGGAAAGGATAAGCCAAGACCAAACGCAGGAGATTTCAGCCATAGATTATCT 1440

Qy        492 TyrIleSerAsnPheAspAlaArgArgSerSerSerGlyGlyIleValSerLeuLeuThr 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TATATATCAAATTTTGATGCACGGCGAAGTAGTTCAGGCGGTATTGTTAGTCTTTTAACG 1500

Qy        512 PheGlyTrpAlaHisThrSerMetAspArgAsnAsnArgLeuGluProAspLysIleThr 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTTGGTTGGGCACATACCAGTATGGATCGTAATAATCGTCTTGAACCAGATAAAATTACT 1560

Qy        532 GlnIleAspAlaValLysGlyTrpGlyGlyAsnIleGlyPheValIleProGlyProThr 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CAAATAGATGCAGTTAAAGGTTGGGGGGGAATATCGGGTTTGTCATCCCAGGACCTACT 1620

Qy        552 GlyGlyAsnLeuValLysValSerAspSerTrpHisSerLeuLysValGlnAlaProGln 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGGGGAATTTGGTAAAAGTCAGTGATAGTTGGCATTCACTTAAAGTTCAAGCACCACAA 1680

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACAAACAAGTTATCGTATTCGTTTGCGTTATGCTTGTTTAGTTACCCATGGGGATGCT 1740

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATTTTTGTAGAACACAGCGGCAGTAGTCATATAGTTTCATTTTTTGATTGCTCAAATTCA 1800

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCAGGTCGTCCATCAAACACTCTTCTAGAGAGTGATTTTCGCTATATTGATGTTCCAGGT 1860

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArgThrGlnSerPheGlyThrHis 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ATTTTTACACCATCAATAAATCCCTTAATAAGATATAGAACACAAAGCTTTGGTACCCAC 1920

Qy        652 AlaIleAspLysPheGluPheIleProLeuAsnThrPheProAsnGlnSerLeuGluLys 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCGATAGACAAATTTGAATTTATTCCACTTAACACTTTTCCGAATCAATCATTAGAAAAA 1980

Qy        672 ArgGluGlnGluValAsnAspLeuPheIleAsn 682
              |||||||||||||||||||||||||||||||||
Db       1981 AGAGAACAGGAAGTAAATGATCTATTTATCAAT 2013
```

```
RESULT 3
CQ868322
LOCUS       CQ868322                1986 bp    DNA     linear   PAT 13-SEP-2004
DEFINITION  Sequence 23 from Patent WO2004074462.
ACCESSION   CQ868322
VERSION     CQ868322.1  GI:51998368
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Carozzi,N., Hargiss,T., Koziel,M.G., Duck,N.B. and Carr,B.
  TITLE     Delta-endotoxin genes and methods for their use
  JOURNAL   Patent: WO 2004074462-A 23 02-SEP-2004;
            Athenix Corporation (US)
FEATURES             Location/Qualifiers
     source          1..1986
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
     CDS             1..1986
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAH33956.1"
                     /db_xref="GI:51998369"
                     /translation="MSNCYPRYPLAKDPQMTMRNTNYKEWLNMCDSNTQFIGDISTYS
                     SPEAALSVRDAVLTGINSVGTILSNLGVPLASQSFGIISRLIGILWAGPDPFEALMVL
                     VEELIKKSIDQRVRENALRELEGLQGIMRLYQTRLQAWLVNKNDDNRRALVTQYAIVD
                     NFFEKNMPKFKERNFEILLLPVYAQAANLHLILLRDADYFGAQWQLGDDEIRDNYIRL
                     QGLIREYKDHCITFYNQGLNQFNRSNAQDWVSFNRFRTDMTLTVLDLAILFPNYDPRR
                     YPLAVKTELTREVYTDPVGFTGVLESGGRTYPWYNPNNTFTAMENNARRRPSYTTWL
                     NRIFVYTRTLGNMSDVRNIWGGHTLVENGNDGSEITHNFGKTDSITPIQYFNFANLSV
                     FSIESLARIYLGGTEANNYITSQYGVSRVIFNTSNINNVPGSLRYEVPANLPSQTILS
                     ELPGKDKPRPNAGDFSHRLSYISNFDARRSSSGGIVSLLTFGWAHTSMDRNNRLEPDK
                     ITQIDAVKGWGGNIGFVIPGPTGGNLVKVSDSWHSLKVQAPQRQTSYRIRLRYACLVT
                     HGDAIFVEHSGSSHIVSFFDCSNSSGRPSNTLLESDFRYIDVPGIFTPSINPLIRYRT
                     QSFGTHAIDKFEFIPLNTFPNQSLEKREQEVNDLFIN"
ORIGIN

Alignment Scores:
Pred. No.:              2.61e-251     Length:         1986
Score:                  3488.00       Matches:        661
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            97.0%         Indels:         0
DB:                     6             Gaps:           0

US-10-782-096-2 (1-682) x CQ868322 (1-1986)
```

```
Qy         22 MetSerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsn 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCTAATTGTTATCCAAGGTATCCACTAGCAAAAGATCCACAAATGACTATGCGAAAC 60

Qy         42 ThrAsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIle 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACGAACTATAAAGAATGGCTAAATATGTGTGATTCAAATACACAATTTATTGGTGATATA 120

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCACGTATTCTAGCCCTGAAGCTGCTTTAAGTGTACGAGATGCTGTTTTAACGGGTATT 180

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGTGTAGGGACTATACTTTCGAATTTAGGGGTCCCTTTGGCAAGTCAATCATTTGGA 240

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAspProPheGluAlaLeuMet 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATAATTAGTAGGCTAATAGGTATTTTATGGGCAGGGCCTGATCCATTTGAAGCACTTATG 300

Qy        122 ValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsnAlaLeu 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTCTTGTTGAAGAGCTTATTAAGAAAAGTATAGATCAGCGTGTAAGAGAAAATGCTCTT 360

Qy        142 ArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGlnAlaTrp 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAGAGCTAGAAGGTTTACAGGGAATTATGAGACTATATCAAACTAGACTGCAAGCATGG 420

Qy        162 LeuValAsnLysAsnAspAspAsnArgArgAlaLeuValThrGlnTyrAlaIleValAsp 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTAGTTAACAAGAATGATGACAATCGGAGGGCACTAGTAACGCAGTATGCAATTGTTGAT 480

Qy        182 AsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIleLeuLeuLeu 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACTTTTTCGAAAAGAATATGCCAAAATTCAAGGAAAGAAACTTTGAAATTTTATTGTTA 540
```

```
Qy        202 ProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAlaAspTyrPhe 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCAGTATATGCACAAGCCGCGAATTTGCATTTAATTTTATTAAGAGATGCTGATTATTTT 600

Qy        222 GlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArgLeuGlnGly 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGCACAGTGGCAATTAGGTGATGATGAAATTCGTGATAATTATATCAGACTACAAGGA 660

Qy        242 LeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeuAsnGlnPhe 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGATTAGAGAATATAAAGATCATTGTATAACATTCTATAACCAGGGTTTAAATCAATTT 720

Qy        262 AsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAspMetThrLeu 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATCGCTCAAATGCTCAAGATTGGGTGAGCTTTAATAGGTTTCGTACAGATATGACATTA 780

Qy        282 ThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyrProLeuAla 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACAGTATTAGATCTCGCAATATTATTTCCAAACTATGATCCACGTAGGTATCCATTAGCA 840

Qy        302 ValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThrGlyValLeu 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTAAAAACGGAATTGACTAGGGAAGTTTATACAGATCCAGTAGGGTTTACTGGGGTATTA 900

Qy        322 GluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPheThrAlaMet 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAAGTGGAGGTAGGACTTACCCTTGGTATAATCCTAATAATACAACCTTTACTGCTATG 960

Qy        342 GluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIlePheValTyr 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAAAATAACGCAAGACGACGTCCTTCTTATACCACTTGGCTTAATCGTATTTTTGTATAT 1020

Qy        362 ThrArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGlyGlyHisThrLeuVal 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACAAGGACTCTAGGTAATATGTCTGATGTGAGAAATATTTGGGGAGGGCATACATTAGTT 1080

Qy        382 GluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLysThrAspSerIleThr 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GAAAATGGAAATGATGGTTCTGAAATAACCCATAACTTTGGTAAAACTGATTCTATTACT 1140

Qy        402 ProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIleGluSerLeuAlaArg 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCTATTCAATATTTTAATTTCGCGAACCTTTCTGTTTTCAGTATTGAGTCACTTGCTCGT 1200

Qy        422 IleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThrSerGlnTyrGlyValSerArg 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATATATTTAGGAGGAACAGAGGCTAATAATTATATTACTAGTCAGTATGGAGTCTCGAGA 1260

Qy        442 ValIlePheAsnThrSerAsnIleAsnAsnValProGlySerLeuArgTyrGluValPro 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTATTTTTAATACATCAAATATAAATAATGTACCTGGATCTTTAAGATACGAAGTGCCT 1320

Qy        462 AlaAsnLeuProSerGlnThrIleLeuSerGluLeuProGlyLysAspLysProArgPro 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCTAATCTTCCATCCCAAACTATATTATCAGAATTACCAGGAAAGGATAAGCCAAGACCA 1380

Qy        482 AsnAlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSer 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACGCAGGAGATTTCAGCCATAGATTATCTTATATATCAAATTTTGATGCACGGCGAAGT 1440

Qy        502 SerSerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArg 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGTTCAGGCGGTATTGTTAGTCTTTTAACGTTTGGTTGGGCACATACCAGTATGGATCGT 1500

Qy        522 AsnAsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGly 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AATAATCGTCTTGAACCAGATAAAATTACTCAAATAGATGCAGTTAAAGGTTGGGGGGGG 1560

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AATATCGGGTTTGTCATCCCAGGACCTACTGGGGGGAATTTGGTAAAAGTCAGTGATAGT 1620

Qy        562 TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArgIleArgLeuArg 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGGCATTCACTTAAAGTTCAAGCACCACAAAGACAAACAAGTTATCGTATTCGTTTGCGT 1680

Qy        582 TyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSerGlySerSerHis 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TATGCTTGTTTAGTTACCCATGGGGATGCTATTTTTGTAGAACACAGCGGCAGTAGTCAT 1740

Qy        602 IleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsnThrLeuLeuGlu 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATAGTTTCATTTTTTGATTGCTCAAATTCATCAGGTCGTCCATCAAACACTCTTCTAGAG 1800

Qy        622 SerAspPheArgTyrIleAspValProGlyIlePheThrProSerIleAsnProLeuIle 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGTGATTTTCGCTATATTGATGTTCCAGGTATTTTTACACCATCAATAAATCCCTTAATA 1860

Qy        642 ArgTyrArgThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeu 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AGATATAGAACACAAAGCTTTGGTACCCACGCGATAGACAAATTTGAATTTATTCCACTT 1920

Qy        662 AsnThrPheProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AACACTTTTCCGAATCAATCATTAGAAAAAAGAGAACAGGAAGTAAATGATCTATTTATC 1980

Qy        682 Asn 682
              |||
Db       1981 AAT 1983

RESULT 4
CS130946
LOCUS       CS130946                3621 bp    DNA     linear   PAT 02-AUG-2005
DEFINITION  Sequence 1 from Patent WO2005066349.
ACCESSION   CS130946
VERSION     CS130946.1  GI:71793183
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Abad,A.
  TITLE     Genes encoding proteins with pesticidal activity
  JOURNAL   Patent: WO 2005066349-A 1 21-JUL-2005;
            E.I. Du pont de nemours and company (US); Pioneer Hi-Bred
            International, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
                     /note="Cry1218-1"
     CDS             1..3621
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAJ21051.1"
                     /db_xref="GI:71793184"
                     /translation="MSPNNQNEYEIIDATPSTSVSNDSNRYPFANEPTNALQNMDYKD
                     YLKMSAGNASEYPGSPEVLVSGQDAAKAAIDIVGKLLSGLGVPFVGPIVSLYTQLIDI
                     LWPSGEKSQWEIFMEQVEELINQKIAEYARNKALSELEGLGNNYQLYLTALEEWEENP
                     NGSRALRDVRNRFEILDSLFTQYMPSFRVTNFEVPFLTVYAMAANLHLLLLKDASIFG
                     EEWGWSTTTINNYYDRQMKLTAEYSDHCVKWYETGLAKLKGTSAKQWVDYNQFRREMT
                     LAVLDVVALFPNYDTRTYPMETKAQLTREVYTDPLGAVNVSSIGSWYDKAPSFGVIES
                     SVIRPPHVFDYITGLTVYTQSRSISSARYIRHWAGHQISYHRVSRGSNLQQMYGTNQN
                     LHSTSTFDFTNYDIYKTLSKDAVLLDIVYPGYTYIFFGMPEVEFFMVNQLNNTRKTLK
                     YNPVSKDIIASTRDSELELPPETSDQPNYESYSHRLCHITSIPATGNTTGLVPVFSWT
                     HRSADLNNTIYSDKITQIPAVKCWDNLPFVPVVKGPGHTGGDLLQYNRSTGSVGTLFL
                     ARYGLALEKAGKYRVRLRYATDADIVLHVNDAQIQMPKTMNPGEDLTSKTFKVADAIT
                     TLNLATDSSLALKHNLGEDPNSTLSGIVYVDRIEFIPVDETYEAEQDLEAAKKAVNAL
                     FTNTKDGLRPGVTDYEVNQAANLVECLSDDLYPNEKRLLFDAVREAKRLSEARNLLQD
                     PDFQEINGENGWTASTGIEVIEGDALFKGRYLRLPGAREIDTETYPTYLYQKVEEGVL
                     KPYTRYRLRGFVGSSQGLEIFTIRHQTNRIVKNVPDDLLPDVSPVNSDGSINRCSEQK
                     YVNSRLEVENRSGEAHEFSIPIDTGEIDYNENAGIWVGFKITDPEGYATLGNLELVEE
```

```
          GPLSGDALERLQREEQQWKIQMTRRREETDRRYMASKQAVDRLYADYQDQQLNPDVEI
          TDLTAAQDLIQSIPYVYNEMFPEIPGMNYTKFTELTDRLQQAWSLYDQRNAIPNGDFR
          NGLSNWNATPGVEVQQINHTSVLVIPNWDEQVSQQFTVQPNQRYVLRVTARKEGVGNG
          YVSIRDGGNQTETLTFSASDYDTNGMYNTQVSNTNGYNTNNAYNTQASSTNGYNANNM
          YNTQASNTNGYNTNSVYNDQTGYITKTVTFIPYTDQMWIEMSETEGTFYIESVELIVD
          VE"
ORIGIN

Alignment Scores:
Pred. No.:              5.86e-61       Length:         3621
Score:                  948.00         Matches:        248
Percent Similarity:     50.9%          Conservative:   119
Best Local Similarity:  34.4%          Mismatches:     276
Query Match:            26.4%          Indels:         78
DB:                     6              Gaps:           25

US-10-782-096-2 (1-682) x CS130946 (1-3621)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db        4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db       61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db      283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db      343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db      403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db      463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db      523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db      583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db      643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db      703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db      763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db      823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db      883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db      943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db      988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db     1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db     1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db     1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy      416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db     1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db     1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db     1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy      470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db     1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db     1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db     1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db     1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db     1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy      559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db     1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db     1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy      592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db     1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812
```

```
Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
Db       2053 AAT 2055
```

```
RESULT 5
CS132867
LOCUS       CS132867                3621 bp    DNA     linear   PAT 02-AUG-2005
DEFINITION  Sequence 5 from Patent WO2005063996.
ACCESSION   CS132867
VERSION     CS132867.1  GI:71791818
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Abad,A.
  TITLE     Plant activation of insect toxin
  JOURNAL   Patent: WO 2005063996-A 5 14-JUL-2005;
            E.I. Du pont de nemours and company (US); Pioneer Hi-Bred
            International, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
ORIGIN

Alignment Scores:
Pred. No.:               5.86e-61       Length:         3621
Score:                   948.00         Matches:        248
Percent Similarity:      50.9%          Conservative:   119
Best Local Similarity:   34.4%          Mismatches:     276
Query Match:             26.4%          Indels:         78
DB:                      6              Gaps:           25

US-10-782-096-2 (1-682) x CS132867 (1-3621)
```

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCTTCAATTGGTTCC-----------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326
```

```
Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||            |||    |||:::   |||
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT-----------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                     |||||||||   :::       :::||||||
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                 :::||||||||||||   :::|||::::::          |||       ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                |||   |||   |||   |||   |||   |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::      |||||||||||||||||   |||||||||   |||      |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||               ||||||   |||||||||:::|||::::::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
              :::      |||      |||:::                        |||   :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::         ||||||:::||||||||||||||               ||||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                    |||   |||            :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::         |||      |||            |||:::      |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                    :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||      ::::::   ||||||   ||||||
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052

Qy        682 Asn 682
              |||
Db       2053 AAT 2055
```

```
RESULT 6
AX543924
LOCUS       AX543924                3621 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO0234774.
ACCESSION   AX543924
VERSION     AX543924.1  GI:25277379
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Abad,A.R., Duck,N.B., Feng,X., Flannagan,R.D., Kahn,T.W. and
            Sims,L.E.
  TITLE     Genes encoding novel proteins with pesticidal activity against
            coleopterans
  JOURNAL   Patent: WO 0234774-A 1 02-MAY-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
                     /note="Cry1218-1"
     CDS             1..3621
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD57542.1"
                     /db_xref="GI:25277380"
                     /translation="MSPNNQNEYEIIDATPSTSVSNDSNRYPFANEPTNALQNMDYKD
                     YLKMSAGNASEYPGSPEVLVSGQDAAKAAIDIVGKLLSGLGVPFVGPIVSLYTQLIDI
                     LWPSGEKSQWEIFMEQVEELINQKIAEYARNKALSELEGLGNNYQLYLTALEEWEENP
                     NGSRALRDVRNRFEILDSLFTQYMPSFRVTNFEVPFLTVYAMAANLHLLLLKDASIFG
                     EEWGWSTTTINNYYDRQMKLTAEYSDHCVKWYETGLAKLKGTSAKQWVDYNQFRREMT
                     LAVLDVVALFPNYDTRTYPMETKAQLTREVYTDPLGAVNVSSIGSWYDKAPSFGVIES
                     SVIRPPHVFDYITGLTVYTQSRSISSARYIRHWAGHQISYHRVSRGSNLQQMYGTNQN
                     LHSTSTFDFTNYDIYKTLSKDAVLLDIVYPGYTYIFFGMPEVEFFMVNQLNNTRKTLK
                     YNPVSKDIIASTRDSELELPPETSDQPNYESYSHRLCHITSIPATGNTTGLVPVFSWT
                     HRSADLNNTIYSDKITQIPAVKCWDNLPFVPVVKGPGHTGGDLLQYNRSTGSVGTLFL
                     ARYGLALEKAGKYRVRLRYATDADIVLHVNDAQIQMPKTMNPGEDLTSKTFKVADAIT
                     TLNLATDSSLALKHNLGEDPNSTLSGIVYVDRIEFIPVDETYEAEQDLEAAKKAVNAL
                     FTNTKDGLRPGVTDYEVNQAANLVECLSDDLYPNEKRLLFDAVREAKRLSEARNLLQD
                     PDFQEINGENGWTASTGIEVIEGDALFKGRYLRLPGAREIDTETYPTYLYQKVEEGVL
                     KPYTRYRLRGFVGSSQGLEIFTIRHQTNRIVKNVPDDLLPDVSPVNSDGSINRCSEQK
                     YVNSRLEVENRSGEAHEFSIPIDTGEIDYNENAGIWVGFKITDPEGYATLGNLELVEE
                     GPLSGDALERLQREEQQWKIQMTRRREETDRRYMASKQAVDRLYADYQDQQLNPDVEI
                     TDLTAAQDLIQSIPYVYNEMFPEIPGMNYTKFTELTDRLQQAWSLYDQRNAIPNGDFR
                     NGLSNWNATPGVEVQQINHTSVLVIPNWDEQVSQQFTVQPNQRYVLRVTARKEGVGNG
                     YVSIRDGGNQTETLTFSASDYDTNGMYNTQVSNTNGYNTNNAYNTQASSTNGYNANNM
                     YNTQASNTNGYNTNSVYNDQTGYITKTVTFIPYTDQMWIEMSETEGTFYIESVELIVD
                     VE"
ORIGIN

Alignment Scores:
Pred. No.:                5.86e-61      Length:          3621
Score:                    948.00        Matches:         248
Percent Similarity:       50.9%         Conservative:    119
Best Local Similarity:    34.4%         Mismatches:      276
Query Match:              26.4%         Indels:          78
DB:                       6             Gaps:            25

US-10-782-096-2 (1-682) x AX543924 (1-3621)
```

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||      |||:::||||||||||||::::::||||||            ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||            |||||||||   |||   :::|||         ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                              |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                |||          |||||||||         :::|||   :::||||||            |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::   ||||||   :::||||||   ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::   ::::::||||||   |||||||||            :::            :::|||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      |||||||||||||||      :::   |||   :::      |||
```

```
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402
Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462
Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522
Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582
Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642
Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702
Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762
Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822
Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882
Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942
Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987
Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1047
Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104
Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164
Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224
Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275
Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326
Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374
Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434
Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494
Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545
Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602
Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662
Qy        559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722
Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767
Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812
Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872
Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932
Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 1992
Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
Db       1993 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2052
Qy        682 Asn 682
Db       2053 AAT 2055

RESULT 7
CS130962
LOCUS       CS130962              4874 bp    DNA     linear   PAT 02-AUG-2005
DEFINITION  Sequence 17 from Patent WO2005066349.
ACCESSION   CS130962
VERSION     CS130962.1  GI:71793198
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Abad,A.
  TITLE     Genes encoding proteins with pesticidal activity
  JOURNAL   Patent: WO 2005066349-A 17 21-JUL-2005;
            E.I. Du pont de nemours and company (US); Pioneer Hi-Bred
            International, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..4874
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
                     /note="Genomic DNA 1218-1"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                8.47e-61      Length:         4874
Score:                    948.00        Matches:        248
Percent Similarity:       50.9%         Conservative:   119
Best Local Similarity:    34.4%         Mismatches:     276
Query Match:              26.4%         Indels:         78
DB:                       6             Gaps:           25

US-10-782-096-2 (1-682) x CS130962 (1-4874)

Qy        3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db      734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy       23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db      791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy       43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db      851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy       63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db      893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy       82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db      953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy      102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db     1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy      120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db     1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy      140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db     1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy      160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db     1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy      179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db     1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy      199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db     1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy      219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db     1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy      239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db     1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy      259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db     1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy      279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db     1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy      299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db     1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy      319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db     1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy      339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db     1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy      359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db     1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy      377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db     1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy      396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db     1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy      416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db     1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy      431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db     2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy      450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db     2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy      470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db     2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy      483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db     2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy      503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db     2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy      523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db     2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy      543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db     2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy      559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
Db     2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy      572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db     2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy      592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db     2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy      612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db     2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy      632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db     2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662
```

```
Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                 :::         |||         :::|||:::    ||||||||||||::::::
Db       2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy        663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
              |||:::         |||   ||||||      ::::::    ||||||   ||||||
Db       2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy        682 Asn 682
              |||
Db       2783 AAT 2785

RESULT 8
AX543950
LOCUS       AX543950                4874 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 27 from Patent WO0234774.
ACCESSION   AX543950
VERSION     AX543950.1  GI:25277423
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Abad,A.R., Duck,N.B., Feng,X., Flannagan,R.D., Kahn,T.W. and
            Sims,L.E.
  TITLE     Genes encoding novel proteins with pesticidal activity against
            coleopterans
  JOURNAL   Patent: WO 0234774-A 27 02-MAY-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1..4874
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
                     /note="Genomic DNA 1218-1"
ORIGIN

Alignment Scores:
Pred. No.:              8.47e-61        Length:         4874
Score:                  948.00          Matches:        248
Percent Similarity:     50.9%           Conservative:   119
Best Local Similarity:  34.4%           Mismatches:     276
Query Match:            26.4%           Indels:         78
DB:                     6               Gaps:           25

US-10-782-096-2 (1-682) x AX543950 (1-4874)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||         |||:::||||||||||||::::::||||||            ::::::::::::
Db        734 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 790

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||            |||||||||   |||   :::|||         ::::::|||
Db        791 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 850

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||:::::::|||   |||                         |||:::   |||
Db        851 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 892

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||         :::|||   :::||||||               |||
Db        893 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 952

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::      ||||||   :::||||||   ||||||||||||
Db        953 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 1012

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
              :::      ::::::|||||||   |||||||||         :::            :::|||
Db       1013 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 1072

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      ||||||||||||||||      :::   |||   :::      |||
Db       1073 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 1132

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||||||||||||            :::||||||   |||   |||:::
Db       1133 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 1192

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||      |||   |||                  |||   :::      ::::::
Db       1193 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 1252

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::   |||   :::   ||||||   |||:::         |||||||||:::
Db       1253 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 1312

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                    |||   |||||||||   ||||||||||||||||||:::||||||:::||||||
Db       1313 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 1372

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    ||||||   :::|||                  |||   :::   |||   |||
Db       1373 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 1432

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                    |||      ||||||   |||||||||:::   :::|||         ||||||
Db       1433 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 1492

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
                 :::         ::::::|||:::      ||||||   :::|||:::||||||   :::
Db       1493 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 1552

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   |||||||||:::      ||||||||||||||||||   |||   |||
Db       1553 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 1612

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||      :::||||||||||||||||||||||||||||:::|||
Db       1613 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 1672

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||         |||                  ||||||:::            :::|||
Db       1673 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 1717

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                 :::|||::::::      |||   |||               ::::::      :::
Db       1718 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTTGATTATATAACGGGACTC 1777

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
              ||||||||||         |||::::::   :::            :::|||:::   |||
Db       1778 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1834

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::                  :::   ||||||   :::      :::|||
Db       1835 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1894

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                    ::::::                  |||:::|||   |||         ::::::
Db       1895 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1954

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                 |||                  |||      :::|||   |||   |||
Db       1955 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 2005

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||            :::|||:::         |||      |||      :::   :::
Db       2006 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 2056

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||                  |||      |||:::   |||
Db       2057 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 2104

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
```

```
                 ::::::                          |||||||||  :::          :::||||||
Db        2105 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 2164

Qy         483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                          :::|||||||||||   :::|||::::::      |||      ::::::
Db        2165 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 2224

Qy         503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
               :::|||   :::                |||  |||  |||  |||  |||  |||
Db        2225 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 2275

Qy         523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
               |||   :::          ||||||||||||||||||   ||||||||  |||      |||
Db        2276 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 2332

Qy         543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
               :::  ||||||                ||||||  |||||||||:::|||::::::
Db        2333 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 2392

Qy         559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
               :::    |||     |||:::                              |||   :::
Db        2393 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 2452

Qy         572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
               :::             ||||||:::||||||||||||||                ||||||
Db        2453 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 2497

Qy         592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                  |||  |||         :::  |||
Db        2498 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 2542

Qy         612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
               :::      |||       |||          |||:::      |||
Db        2543 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 2602

Qy         632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
                           :::   :::      :::    :::
Db        2603 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 2662

Qy         645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn--- 662
                  :::         |||         :::|||:::   ||||||||||||::::::
Db        2663 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGATGAG 2722

Qy         663 ThrPhe---ProAsnGlnSerLeuGluLysArgGluGlnGluValAsnAspLeuPheIle 681
               |||:::       |||   ||||||        ::::::    ||||||  ||||||
Db        2723 ACATATGAAGCGGAACAAGATTTAGAAGCAGCGAAGAAAGCAGTGAATGCCTTGTTTACG 2782

Qy         682 Asn 682
               |||
Db        2783 AAT 2785
```

```
RESULT 9
BTU04365
LOCUS       BTU04365                3507 bp    DNA     linear   BCT 27-AUG-1994
DEFINITION  Bacillus thuringiensis kumamotoensis PS50C(b) CryIII
            delta-endotoxin  gene, partial cds.
ACCESSION   U04365
VERSION     U04365.1  GI:436834
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1  (bases 1 to 3507)
  AUTHORS   Narva,K.E. and Fu,J.
  TITLE     Novel Coleopteran-Active Toxins from Bacillus thuringiensis
  JOURNAL   Unpublished (1994)
REFERENCE   2  (bases 1 to 3507)
  AUTHORS   Feitelson,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-1993) Jerald S. Feitelson, Molecular Biology,
            Mycogen Corporation, 4980 Carroll Canyon Road, San Diego, CA 92121,
            USA
FEATURES             Location/Qualifiers
     source          1..3507
                     /organism="Bacillus thuringiensis"
                     /mol_type="genomic DNA"
                     /strain="PS50C(b)"
                     /sub_species="kumamotoensis"
                     /db_xref="taxon:1428"
                     /note="serotype 18"
     CDS             1..>3507
                     /codon_start=1
                     /transl_table=11
                     /product="CryIII delta-endotoxin"
                     /protein_id="AAA21118.1"
                     /db_xref="GI:436835"
                     /translation="MSPNNQNEYEIIDATPSTSVSNDSNRYPFANEPTNALQNMDYKD
                     YLKMSAGNVSEYPGSPEVFLSEQDAVKAAIDIVGKLLTGLGVPFVGPIVSLYTQLIDI
                     LWPSKQKSQWEIFMEQVEELINQKIAEYARNKALSELEGLGNNYQLYLTALEEWKENP
                     NGSRALRDVRNRFEILDSLFTQYMPSFRVTNFEVPFLTVYTMAANLHLLLLRDASIFG
                     EEWGLSTSTINNYYNRQMKLTAEYSDHCVKWYETGLAKLKGSSAKQWIDYNQFRREMT
                     LTVLDVVALFSNYDTRTYPLATTAQLTREVYTDPLGAVDVPNIGSWYDKAPSFSEIEK
                     AAIRPPHVFDYITGLTVYTKKRSFTSDRYMRYWAGHQISYKHIGTSSTFTQMYGTNQN
                     LQSTSNFDFTNYDIYKTLSNGAVLLDIVYPGYTYTFFGMPETEFFMVNQLNNTRKTLT
                     YKPASKDIIDRTRDSELELPPETSGQPNYESYSHRLGHITFIYSSSTSTYVPVFSWTH
                     RSADLTNTVKSGEITQIPGGKSSTIGRNTYIIKGRGYTGGDLVALTDRIGSCEFQMIF
                     PESQRFRIRIRYASNETSYISLYGLNQSGTLKFNQTYSNKNENDLTYNDFKYIEYPRV
                     ISVNASSNIQRLSIGIQTNTNLFILDRIEFIPVDETYEAETDLEAAKKAVNALFTNTK
                     DGLQPGVTDYEVNQAANLVECLSDDLYPNEKRLLFDAVREAKRLSEARNLLQDPDFQE
                     INGENGWTASTGIEVIEGDAVFKGRYLRLPGAREIDTETYPTYLYQKVEEGVLKPYTR
                     YRLRGFVGSSQGLEIYTIRHQTNRIVKNVPDDLLPDVPPVNNDGRINRCSEQKYVNSR
                     LEVENRSGEAHEFSIPIDTGELDYNENAGIWVGFKITDPEGYATLGNLELVEEGPLSG
                     DALERLQKEEQQWKIQMTRRREETDRRYMASKQAVDRLYADYQDQQLNPNVEITDLTA
                     AQDLIQSIPYVYNEMFPEIPGMNYTKFTELTDRLQQAWGLYDQRNAIPNGDYRNELSN
                     WNTTSGVNVQQINHTSVLVIPNWNEQVSQKFTVQPNQRYVLRVTARKEGVGNGYVSIR
                     DGGNQSETLTFSASDYDTNGMYDTQASNTNGYNTNSVYMIKPAISRKTVDISSVYNQM
                     WIEISETEGTFYIESVELIVDVE"
ORIGIN

Alignment Scores:
Pred. No.:              1.45e-60        Length:         3507
Score:                  942.50          Matches:        249
Percent Similarity:     50.3%           Conservative:   115
Best Local Similarity:  34.4%           Mismatches:     272
Query Match:            26.2%           Indels:         87
DB:                     1               Gaps:           27

US-10-782-096-2 (1-682) x BTU04365 (1-3507)

Qy           3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                |||      |||:::||||||||||||::::::||||||           ::::::::::::
Db           4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy          23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
               ||||||         |||||||||  |||    :::|||          ::::::|||
Db          61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy          43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
               :::||||||::::::|||   |||                      |||::::::|||
Db         121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGTTAGT 162

Qy          63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                  |||        |||||||||         |||      :::||||||||||        |||
Db         163 GAATACCCTGGTTCACCTGAGGTATTTCTAAGCGAGCAAGATGCAGTTAAGGCCGCAATT 222

Qy          82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
               :::   ||||||   :::|||:::   ||||||||||||
Db         223 GATATAGTAGGTAAATTACTAACAGGTTTAGGGGTTCCATTTGTTGGGCCGATAGTGAGT 282

Qy         102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGly------ProAspProPheGluAla 119
               :::   ::::::||||||   |||||||||                        :::|||
Db         283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAAAACAAAAGAGTCAATGGGAAATT 342

Qy         120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                  |||      ||||||||||||||   :::   |||   :::      |||
```

```
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        403 GCGCTTTCGGAATTGGAAGGGCTAGGGAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db        463 GAGTGGAAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTTCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTCGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        583 CCATTCCTTACAGTATATACAATGGCAGCAAACCTACATTTACTTTTATTAAGGGACGCA 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        643 TCAATTTTTGGAGAAGAATGGGGATTGTCTACAAGCACTATTAATAACTACTATAATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        703 CAAATGAAACTTACTGCAGAATATTCTGACCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        763 GCAAAATTAAAAGGCTCGAGCGCTAAACAATGGATTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        823 ATGACATTGACGGTGTTAGACGTTGTTGCATTATTTTCAAACTATGATACGCGTACGTAT 882

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        883 CCACTGGCAACAACAGCTCAGCTTACAAGGGAAGTATATACAGATCCACTTGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        943 GATGTGCCTAATATTGGCTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        988 TCAGAAATAGAAAAAGCGGCTATTCGTCCACCTCATGTGTTTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThrArgThrLeuGlyAsnMetSerAsp---ValArgAsnIleTrpGlyGly 377
Db       1048 ACAGTTTATACAAAAAAACGTAGCTTCACTTCTGATCGTTATATGAGATATTGGGCTGGT 1107

Qy        378 HisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLys 396
Db       1108 CATCAAATAAGCTATAAGCATATCGGTACGAGTAGTACCTTTACACAGATGTATGGAACC 1167

Qy        397 ThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIle 416
Db       1168 AATCAAAATTTACAAAGTACTAGCAATTTTGATTTTACGAATTACGATATTTACAAGACT 1227

Qy        417 GluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsn 431
Db       1228 TTATCAAATGGTGCAGTACTCCTTGATATAGTTTACCCTGGT------------------ 1269

Qy        432 TyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIleAsn 450
Db       1270 TATACGTATACATTTTTTGGAATGCCAGAAACCGAGTTTTTTATGGTAAATCAATTGAAT 1329

Qy        451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
Db       1330 AATACCAGAAAGACGTTAACGTATAAA---CCAGCT---------TCCAAAGATATTATA 1377

Qy        471 Ser--------------------GluLeuProGlyLysAspLysProArgProAsnAla 483
Db       1378 GATCGGACAAGAGATTCGGAATTAGAATTGCCTCCAGAAACTTCAGGTCAACCAAATTAC 1437

Qy        484 GlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSer 503
Db       1438 GAGTCATATAGCCATAGATTAGGTCATATTACATTTATT-----------TACTCCAGT 1485

Qy        504 GlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsn 523
Db       1486 TCAACTAGCACGTATGTACCTGTATTTTCTTGGACACATCGGAGTGCAGATCTAACAAAT 1545

Qy        524 ArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIle 543
Db       1546 ACAGTTAAAAGTGGCGAAATCACCCAAATACCAGGG------GGCAAGTCTAGCACCATA 1599

Qy        544 Gly---------PheValIleProGlyPro------ThrGlyGlyAsnLeuValLysVal 558
Db       1600 GGCAGAAATACTTATATAATAAAAGGGCGTGGTTATACAGGGGGAGACTTAGTGGCTTTA 1659

Qy        559 SerAspSerTrpHisSerLeuLysValGlnAla------ProGlnArgGlnThrSerTyr 576
Db       1660 ACGGACCGCATCGGAAGTTGTGAGTTTCAGATGATCTTTCCAGAGTCTCAA---CGATTC 1716

Qy        577 ArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHis 596
Db       1717 CGTATTCGGATTCGTTACGCT--------------------------------------- 1737

Qy        597 SerGlySerSerHisIleValSerPhePheAspCysSerAsnSer--------------- 611
Db       1738 TCTAATGAAACTAGTTATATTAGTTTATACGGACTAAACCAAAGCGGAACTTTAAAATTC 1797

Qy        612 ------------SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIle 627
Db       1798 AACCAGACATATTCTAATAAAAATGAAAATGATTTAACATATAATGATTTCAAATATATA 1857

Qy        628 AspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArgThrGlnSer 647
Db       1858 GAATATCCAAGAGTCATT------TCAGTAAATGCTTCTTCAAACATACAGAGGTTATCT 1911

Qy        648 PheGlyThrHisAla------------------IleAspLysPheGluPheIleProLeu 661
Db       1912 ATAGGTATACAAACGAATACAAATTTATTATTTTAGACCGAATCGAATTCATCCCAGTA 1971

Qy        662 Asn---ThrPheProAsnGlnSer---LeuGluLysArgGluGlnGluValAsnAspLeu 679
Db       1972 GATGAGACATATGAAGCGGAAACGGATTTAGAAGCGGCAAAGAAAGCAGTGAATGCCTTG 2031

Qy        680 PheIleAsn 682
Db       2032 TTTACGAAT 2040

RESULT 10
I25972
LOCUS       I25972                  3507 bp    DNA     linear   PAT 07-OCT-1996
DEFINITION  Sequence 3 from patent US 5554534.
ACCESSION   I25972
VERSION     I25972.1  GI:1605842
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3507)
  AUTHORS   Michaels,T.E., Narva,K.E. and Foncerrada,L.
  TITLE     Bacillus thuringiensis toxins active against scarab pests
  JOURNAL   Patent: US 5554534-A 3 10-SEP-1996;
FEATURES             Location/Qualifiers
     source          1..3507
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:            1.45e-60      Length:         3507
Score:                942.50        Matches:        249
Percent Similarity:   50.3%         Conservative:   115
```

```
Best Local Similarity:  34.4%     Mismatches:  272
Query Match:            26.2%     Indels:       87
DB:                     6         Gaps:         27

US-10-782-096-2 (1-682) x I25972 (1-3507)

Qy      3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db      4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACATCTGTA 60

Qy     23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db     61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy     43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
Db    121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGTTAGT 162

Qy     63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db    163 GAATACCCTGGTTCACCTGAGGTATTTCTAAGCGAGCAAGATGCAGTTAAGGCCGCAATT 222

Qy     82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db    223 GATATAGTAGGTAAATTACTAACAGGTTTAGGGGTTCCATTTGTTGGGCCGATAGTGAGT 282

Qy    102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGly------ProAspProPheGluAla 119
Db    283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAAAACAAAAGAGTCAATGGGAAATT 342

Qy    120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db    343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy    140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db    403 GCGCTTTCGGAATTGGAAGGGCTAGGGAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy    160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db    463 GAGTGGAAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTTCGAAATCGATTTGAA 522

Qy    179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db    523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTCGAGTGACAAATTTTGAAGTA 582

Qy    199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db    583 CCATTCCTTACAGTATATACAATGGCAGCAAACCTACATTTACTTTTATTAAGGGACGCA 642

Qy    219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db    643 TCAATTTTTGGAGAAGAATGGGGATTGTCTACAAGCACTATTAATAACTACTATAATGGT 702

Qy    239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db    703 CAAATGAAACTTACTGCAGAATATTCTGACCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db    763 GCAAAATTAAAAGGCTCGAGCGCTAAACAATGGATTGACTATAACCAATTCCGTAGAGAA 822

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db    823 ATGACATTGACGGTGTTAGACGTTGTTGCATTATTTTCAAACTATGATACGCGTACGTAT 882

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db    883 CCACTGGCAACAACAGCTCAGCTTACAAGGGAAGTATATACAGATCCACTTGGCGCGGTA 942

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db    943 GATGTGCCTAATATTGGCTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db    988 TCAGAAATAGAAAAAGCGGCTATTCGTCCACCTCATGTGTTTGATTATATAACGGACTC 1047

Qy    359 PheValTyrThrArgThrLeuGlyAsnMetSerAsp---ValArgAsnIleTrpGlyGly 377
Db   1048 ACAGTTTATACAAAAAAACGTAGCTTCACTTCTGATCGTTATATGAGATATTGGGCTGGT 1107

Qy    378 HisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGlyLys 396
Db   1108 CATCAAATAAGCTATAAGCATATCGGTACGAGTAGTACCTTTACACAGATGTATGGAACC 1167

Qy    397 ThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSerIle 416
Db   1168 AATCAAAATTTACAAAGTACTAGCAATTTTGATTTTACGAATTACGATATTTACAAGACT 1227

Qy    417 GluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsn 431
Db   1228 TTATCAAATGGTGCAGTACTCCTTGATATAGTTTACCCTGGT------------------ 1269

Qy    432 TyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIleAsn 450
Db   1270 TATACGTATACATTTTTTGGAATGCCAGAAACCGAGTTTTTTATGGTAAATCAATTGAAT 1329

Qy    451 AsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIleLeu 470
Db   1330 AATACCAGAAAGACGTTAACGTATAAA---CCAGCT---------TCCAAAGATATTATA 1377

Qy    471 Ser---------------------GluLeuProGlyLysAspLysProArgProAsnAla 483
Db   1378 GATCGGACAAGAGATTCGGAATTAGAATTGCCTCCAGAAACTTCAGGTCAACCAAATTAC 1437

Qy    484 GlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSerSer 503
Db   1438 GAGTCATATAGCCATAGATTAGGTCATATTACATTTATT------------TACTCCAGT 1485

Qy    504 GlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsn 523
Db   1486 TCAACTAGCACGTATGTACCTGTATTTTCTTGGACACATCGGAGTGCAGATCTAACAAAT 1545

Qy    524 ArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsnIle 543
Db   1546 ACAGTTAAAAGTGGCGAAATCACCCAAATACCAGGG------GGCAAGTCTAGCACCATA 1599

Qy    544 Gly---------PheValIleProGlyPro------ThrGlyGlyAsnLeuValLysVal 558
Db   1600 GGCAGAAATACTTATATAATAAAAGGGCGTGGTTATACAGGGGGAGACTTAGTGGCTTTA 1659

Qy    559 SerAspSerTrpHisSerLeuLysValGlnAla------ProGlnArgGlnThrSerTyr 576
Db   1660 ACGGACCGCATCGGAAGTTGTGAGTTTCAGATGATCTTTCCAGAGTCTCAA---CGATTC 1716

Qy    577 ArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHis 596
Db   1717 CGTATTCGGATTCGTTACGCT--------------------------------------- 1737

Qy    597 SerGlySerSerHisIleValSerPhePheAspCysSerAsnSer--------------- 611
Db   1738 TCTAATGAAACTAGTTATATTAGTTTATACGGACTAAACCAAAGCGGAACTTTAAAATTC 1797

Qy    612 ------------SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIle 627
Db   1798 AACCAGACATATTCTAATAAAAATGAAAATGATTTAACATATAATGATTTCAAATATATA 1857

Qy    628 AspValProGlyIlePheThrProSerIleAsnProLeuIleArgTyrArgThrGlnSer 647
Db   1858 GAATATCCAAGAGTCATT------TCAGTAAATGCTTCTTCAAACATACAGAGGTTATCT 1911

Qy    648 PheGlyThrHisAla------------------IleAspLysPheGluPheIleProLeu 661
Db   1912 ATAGGTATACAAACGAATACAAATTTATTTATTTAGACCGAATCGAATTCATCCCAGTA 1971
```

```
Qy        662 Asn---ThrPheProAsnGlnSer---LeuGluLysArgGluGlnGluValAsnAspLeu 679
              :::   |||:::        ::::::  ||||||     ::::::  ||||||   |||
Db       1972 GATGAGACATATGAAGCGGAAACGGATTTAGAAGCGGCAAAGAAAGCAGTGAATGCCTTG 2031

Qy        680 PheIleAsn 682
              |||  |||
Db       2032 TTTACGAAT 2040

RESULT 11
BD133574
LOCUS       BD133574               3504 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Protein having insecticidal activity, DNA encoding the protein, and
            controlling agent and controlling method of noxious organisms.
ACCESSION   BD133574
VERSION     BD133574.1  GI:23228519
KEYWORDS    JP 2002045186-A/1.
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1  (bases 1 to 3504)
  AUTHORS   Asano,S., Yamanaka,S. and Takeuchi,K.
  TITLE     Protein having insecticidal activity, DNA encoding the protein, and
            controlling agent and controlling method of noxious organisms
  JOURNAL   Patent: JP 2002045186-A 1 12-FEB-2002;
            SDS BIOTECH KK
COMMENT     OS   Bacillus thuringiensis
            PN   JP 2002045186-A/1
            PD   12-FEB-2002
            PF   03-AUG-2000 JP 2000236140
            PI   SHINICHIRO ASANO,SATOSHI YAMANAKA,KATSUYOSHI TAKEUCHI PC
            C12N15/09,A01N63/00,A01N63/02,C07K14/325,C12N1/15,C12N1/19, PC
            C12N1/20,
            PC   C12N1/20,C12N1/21,C12N5/10,C12N5/10//(C12N15/09,C12R1:19), PC
             (C12N1/20,C12R1:07),C12N15/00,C12N5/00,C12N5/00,(C12N15/00, PC
            C12R1:19)
            CC   Protein having insecticidal activity, DNA
            encoding the protein,
            CC        and
            CC   controlling agent and controlling method of noxious organisms
            FH   Key             Location/Qualifiers
            FT   CDS             (1)..(3501).
FEATURES             Location/Qualifiers
     source          1..3504
                     /organism="Bacillus thuringiensis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1428"
ORIGIN

Alignment Scores:
Pred. No.:              2.05e-60        Length:         3504
Score:                  940.50          Matches:        256
Percent Similarity:     51.8%           Conservative:   111
Best Local Similarity:  36.1%           Mismatches:     288
Query Match:            26.1%           Indels:         56
DB:                     6               Gaps:           23

US-10-782-096-2 (1-682) x BD133574 (1-3504)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||      |||:::|||||||||||:::|||||||||        ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              |||:::         |||||||||||||   |||         |||:::::::|||
Db         61 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61
              ||||||||::::::|||   |||   :::      |||   :::    |||:::
Db        121 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||:::   |||                              |||   |||||||||
Db        181 GAGACGTTTATTAGT--------------------TCATCTACGGTTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
                       ||||||   :::|||     |||||||||||   |||   |||
Db        220 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
                       |||   ::::::|||   ||||||                   :::|||
Db        280 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
              :::|||      ||||||:::||||||   :::      |||         ||||||:::
Db        340 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||      |||:::||||||         :::   :::||||||      |||:::
Db        400 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 459

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178
                 ||||||   |||:::||||||            |||:::   :::|||||||||||
Db        460 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 519

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
                 ::::::         |||         :::|||   |||            |||:::
Db        520 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 579

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                 ||||||   |||||||||||||||||||||||||:::||||||||||||||||
Db        580 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 639

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                    |||||||||:::|||                  ||||||         |||   |||
Db        640 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 699

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                             :::|||   |||:::|||:::   :::||||||   ||||||
Db        700 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 759

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
              ::::::         :::||||||      |||:::   :::::::::||||||   :::
Db        760 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 819

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   ||||||||||||      ||||||||||||||||||   |||   |||
Db        820 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 879

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::            :::||||||||||||||||||||||||:::      |||
Db        880 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 939

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                          |||||||||         |||            ::::::   :::|||
Db        940 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTTCTTTT 996

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
              :::   :::|||:::            |||   |||            |||:::   |||
Db        997 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTTGATATACTCAGTGAAATA 1056

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375
                 ||||||         |||                 ::::::            |||
Db       1057 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1116

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
                 ||||||::::::      :::|||   |||      |||   :::      |||:::
Db       1117 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1176

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
              |||   |||   :::         |||:::|||::::::   ||||||      :::|||
Db       1177 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1233

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
                 :::   ||||||                     |||   :::      ||||||
Db       1234 CAGGTTCGATCATTA-----------------GGGGCGGATTTAGCTAATTACTACGCA 1275
```

```
Qy     435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
Db    1276 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1335

Qy     455 SerLeu---------ArgTyrGluValProAlaAsn---------------LeuProSer 466
Db    1336 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1395

Qy     467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
Db    1396 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1449

Qy     487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504
Db    1450 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTTCTAAGAATGCTAGTAGTCCTGCT 1509

Qy     505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524
Db    1510 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1569

Qy     525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541
Db    1570 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1629

Qy     542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
Db    1630 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1689

Qy     562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577
Db    1690 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1746

Qy     578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597
Db    1747 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1800

Qy     598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617
Db    1801 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGAAT 1851

Qy     618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636
Db    1852 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 1911

Qy     637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653
Db    1912 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 1971

Qy     654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671
Db    1972 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2031

Qy     672 ArgGluGlnGluValAsnAspLeuPhe 680
Db    2032 GCAAAGAAAGCAGTAAATGGCTTGTTT 2058
```

```
RESULT 12
BD133575
LOCUS       BD133575                3690 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Protein having insecticidal activity, DNA encoding the protein, and
            controlling agent and controlling method of noxious organisms.
ACCESSION   BD133575
VERSION     BD133575.1  GI:23228520
KEYWORDS    JP 2002045186-A/2.
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1  (bases 1 to 3690)
  AUTHORS   Asano,S., Yamanaka,S. and Takeuchi,K.
  TITLE     Protein having insecticidal activity, DNA encoding the protein, and
            controlling agent and controlling method of noxious organisms
  JOURNAL   Patent: JP 2002045186-A 2 12-FEB-2002;
            SDS BIOTECH KK
COMMENT     OS   Bacillus thuringiensis
            PN   JP 2002045186-A/2
            PD   12-FEB-2002
            PF   03-AUG-2000 JP 2000236140
            PI   SHINICHIRO ASANO,SATOSHI YAMANAKA,KATSUYOSHI TAKEUCHI PC
            C12N15/09,A01N63/00,A01N63/02,C07K14/325,C12N1/15,C12N1/19, PC
            C12N1/20,
            PC   C12N1/20,C12N1/21,C12N5/10,C12N5/10//(C12N15/09,C12R1:19), PC
             (C12N1/20,C12R1:07),C12N15/00,C12N5/00,C12N5/00,(C12N15/00, PC
            C12R1:19)
            CC   Protein having insecticidal activity, DNA
            encoding the protein,
            CC          and
            CC   controlling agent and controlling method of noxious organisms
            FH   Key             Location/Qualifiers
            FT   source          1. .3690
            FT                   /organism='Bacillus thuringiensis'.
FEATURES             Location/Qualifiers
     source          1. .3690
                     /organism="Bacillus thuringiensis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1428"
ORIGIN

Alignment Scores:
Pred. No.:              2.18e-60      Length:         3690
Score:                  940.50        Matches:        256
Percent Similarity:     51.8%         Conservative:   111
Best Local Similarity:  36.1%         Mismatches:     288
Query Match:            26.1%         Indels:         56
DB:                     6             Gaps:           23

US-10-782-096-2 (1-682) x BD133575 (1-3690)

Qy       3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db     190 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 246

Qy      23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db     247 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 306

Qy      43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61
Db     307 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 366

Qy      62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db     367 GAGACGTTTATTAGT---------------------TCATCTACGGTTCAAACTGGAATT 405

Qy      82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db     406 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 465

Qy     102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
Db     466 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 525

Qy     120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db     526 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 585

Qy     140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db     586 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 645

Qy     160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178
Db     646 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 705

Qy     179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
```

```
                    ::::::       |||          :::|||   |||                    |||:::
Db        706 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 765

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
              ||||||  ||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        766 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 825

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                       |||||||||:::|||          ||||||       |||    |||
Db        826 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 885

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                                 :::|||  |||:::|||:::   :::|||||  ||||||
Db        886 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 945

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
              ::::::        :::||||||      |||:::   :::::::::||||||   :::
Db        946 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 1005

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||  ||||||||||||      ||||||||||||||||||||   |||   |||
Db       1006 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 1065

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::           :::||||||||||||||||||||||||||||:::   |||
Db       1066 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 1125

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                          |||||||||       |||          ::::::   :::|||
Db       1126 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTTCTTTT 1182

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
              :::   :::|||:::          |||   |||          |||:::   |||
Db       1183 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTTGATATACTCAGTGAAATA 1242

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375
                    ||||||      |||          |||   ::::::            |||
Db       1243 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1302

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
                 ||||||::::::          :::|||   |||      |||   :::   |||:::
Db       1303 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1362

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
              |||   |||   :::         |||:::||||:::::      ||||||     :::|||
Db       1363 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1419

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
                 :::   ||||||                  |||   :::      ||||||
Db       1420 CAGGTTCGATCATTA------------------GGGGCGGATTTAGCTAATTACTACGCA 1461

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
                    |||||||||          |||         :::   |||            |||
Db       1462 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1521

Qy        455 SerLeu--------ArgTyrGluValProAlaAsn--------------LeuProSer 466
              |||:::                |||   |||                         :::
Db       1522 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1581

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
                 :::      |||:::|||   :::::::::|||                   :::
Db       1582 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1635

Qy        487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504
              |||||||||||||||:::|||:::         :::|||         ||||||
Db       1636 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTCTAAGAATGCTAGTAGTCCTGCT 1695

Qy        505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524
                       :::|||      |||   |||   |||   |||   |||        |||
Db       1696 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1755

Qy        525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541
              :::         ||||||||||||||||||       ||||||              |||
Db       1756 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1815

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
                                ||||||   ||||||||||||::::::|||   :::   |||
Db       1816 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1875

Qy        562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577
                           :::   |||:::   |||::::::|||       |||      ||||||
Db       1876 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1932

Qy        578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597
                 |||:::|||||||||         |||             :::||||||      |||
Db       1933 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1986

Qy        598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617
                 |||   ||||||::::::      :::   |||   |||   :::           :::
Db       1987 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGGAT 2037

Qy        618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636
                 |||      :::   |||            :::      ||||||         :::
Db       2038 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 2097

Qy        637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653
                 :::      :::                  ||||||         |||         :::
Db       2098 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 2157

Qy        654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671
              ||||||||||||   |||||||:::|||   ||||||         :::   |||:::
Db       2158 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2217

Qy        672 ArgGluGlnGluValAsnAspLeuPhe 680
                 ::::::   ||||||   ||||||
Db       2218 GCAAAGAAAGCAGTAAATGGCTTGTTT 2244
```

```
RESULT 13
AB089299
LOCUS       AB089299                3435 bp    DNA     linear   BCT 20-DEC-2003
DEFINITION  Bacillus thuringiensis serovar galleriae cry8 gene, complete cds.
ACCESSION   AB089299
VERSION     AB089299.1  GI:22122187
KEYWORDS    .
SOURCE      Bacillus thuringiensis serovar galleriae
  ORGANISM  Bacillus thuringiensis serovar galleriae
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Asano,S., Yamashita,C., Iizuka,T., Takeuchi,K., Yamanaka,S.,
            Cerf,D. and Yamamoto,T.
  TITLE     A strain of Bacillus thuringiensis subsp. galleriae containing a
            novel cry8 gene highly toxic to Anomala cuprea (Coleoptera:
            Scarabaeidae)
  JOURNAL   Biol. Control 28, 191-196 (2003)
REFERENCE   2  (bases 1 to 3435)
  AUTHORS   Asano,S. and Yamamoto,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2002) Shin-ichiro Asano, Hokkaido University,
            Graduate school of Agriculture; N9W9, Sapporo, Hokkaio 060-8589,
            Japan (E-mail:sangaku@abs.agr.hokudai.ac.jp, Tel:81-11-706-2423,
            Fax:81-11-706-2423)
FEATURES             Location/Qualifiers
     source          1..3435
                     /organism="Bacillus thuringiensis serovar galleriae"
                     /mol_type="genomic DNA"
                     /strain="SDS-502"
                     /serovar="galleriae"
                     /db_xref="taxon:29338"
     gene            1..3435
                     /gene="cry8"
     CDS             1..3435
                     /gene="cry8"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC07226.1"
```

```
                      /db_xref="GI:22122188"
                      /translation="MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNMNYKD
                      YLRMSEGENPELFGNPETFISSSTVQTGIGIVGQVLGALGVPFAGQIASFYSFIVGQL
                      WPSSTVSVWEMIMKQVEDLIDQKITDSVRKTALAGLQGLGDGLDVYQKSLKNWLENRN
                      DTRARSVVVTQYIALELDFVAKIPSFAISGQEVPLLSVYAQAANLHLLLLRDASIFGA
                      EWGFTPGEISTFYDRQVTRTAQYSDYCVKWYNTGLDKLKGTNAASWLKYHQFRREMTL
                      LVLDLVALFPNYDTRTYPIETTAQLTREVYTDPIVFNRETSGGFCRRWSLNSDISFSE
                      VESAVIRSPHLFDILSEIEFYTTRAGLPLNNTEYLEYWVGHSIKYKNTNASSALERNY
                      GTITSNKIKYYDLANKDIFQVRSLGADLANYYAQVYGVPYASFTLLDKNTGSGSVGGF
                      TYSKPHTTMQVCTQNYNTIDEIPPENEPLSRGYSHRLSHITSYSFSKNASSPARYGNL
                      PVFAWTHRSADVTNTVYSDKITQIPVVKAHTLVSGTTVIKGPGFTGGNILKRTSSGPL
                      AYTSVSVKSPLSQRYRARIRYASTTNLRLFVTISGTRIYSINVNKTMNKGDDLTFNTF
                      DLATIGTAFTFSNYSDSLTVGADSFASGGEVYVDKFELIPVNATFEAEEDLDVAKKAV
                      KNLVECLSDELYPNEKRMLWDAVKEAKRLVQARNLLQDTGFNRINGENGWTGSTGIEV
                      AEGDVLFKDRSLRLTSAREIDTETYPTYLYQQIDESLLKPYTRYKLKGFIGSSQDLEI
                      KLIRHRANQIVKNVPDNLLPDVLPVNSCGGIDRCSEQQYVDANLALENNGENGNMSSD
                      SHAFSFHIDTGEIDLNENTGIWVVFKIPTTNGYATLGNLELVEEGPLSGETLERAQQQ
                      EQQWQDKMARKRGASEKAYYAAKQAIDRLFADYQDQKLNSGVEMSDMLAAQNLVQSIP
                      YVYNDALPEIPGMNYTSFTELTNRLQQAWNLYDLRNAIPNGDFRNGLSDWNATSDVNV
                      QQLSDTSVLVIPNWNSQVSQQFTVQPNYRYVLRVTARKEGVGDGYVIIRDGANQTETL
                      TFNICDDDTGVLSADQTSYITKTVEFTPSTEQVWIDMSETEGVFNIESVELVLEEE"
ORIGIN

Alignment Scores:
Pred. No.:              1.12e-59       Length:         3435
Score:                  930.50         Matches:        254
Percent Similarity:     51.7%          Conservative:   112
Best Local Similarity:  35.9%          Mismatches:     288
Query Match:            25.9%          Indels:         56
DB:                     1              Gaps:           23

US-10-782-096-2 (1-682) x AB089299 (1-3435)

Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTCTAGATGCT---TCATCATCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
Db         61 TCCGATAATTCTGTTAGATACCCTTTAGCAAACGATCAAACGACCACATTACAAAACATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAsp---SerAsnThrGlnPheIleGlyAspIle 61
Db        121 AACTATAAAGATTATCTGAGAATGTCTGAGGGAGAGAATCCTGAATTATTTGGAAATCCG 180

Qy         62 SerThrTyrSerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
Db        181 GAGACGTTTATTAGT--------------------TCATCTACGGTTCAAACTGGAATT 219

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db        220 GGCATTGTTGGTCAAGTACTGGGGGCTTTAGGGGTTCCATTTGCTGGACAGATAGCTAGT 279

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyPro------AspProPheGluAla 119
Db        280 TTTTATAGTTTCATTGTCGGTCAATTATGGCCATCAAGTACCGTGAGTGTATGGGAAATG 339

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db        340 ATTATGAAACAAGTGGAAGATCTAATTGATCAAAAAATAACAGATTCTGTAAGGAAAACA 399

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db        400 GCGCTTGCAGGACTACAAGGATTAGGAGATGGCTTAGACGTATATCAGAAATCACTTAAG 459

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArgArgAla---LeuValThrGlnTyrAla 178
Db        460 AATTGGCTGGAAAATCGTAATGATACAAGAGCTAGAAGTGTTGTGGTGACCCAATATATA 519

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db        520 GCTTTAGAGCTTGATTTTGTTGCTAAAATCCCATCTTTTGCAATATCTGGACAGGAAGTA 579

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db        580 CCATTATTATCAGTGTATGCACAAGCAGCGAATTTACATTTGCTATTATTACGAGATGCT 639

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db        640 TCCATTTTTGGAGCAGAGTGGGGATTCACACCAGGAGAAATTTCCACATTTTATGATCGT 699

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db        700 CAGGTGACACGTACCGCCCAATACTCGGATTATTGTGTAAAGTGGTATAACACTGGCTTA 759

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db        760 GATAAATTAAAAGGTACGAATGCTGCAAGTTGGCTGAAGTATCACCAATTCCGAAGAGAA 819

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db        820 ATGACATTACTGGTATTAGATTTAGTAGCGTTATTTCCAAACTATGACACACGTACGTAT 879

Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db        880 CCAATCGAAACAACGGCCCAACTTACACGGGAAGTGTATACAGATCCAATAGTATTTAAC 939

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db        940 AGAGAA---ACAAGTGGTGGATTTTGTAGGCGTTGGTCACTTAACAGTGATATTTCTTTT 996

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db        997 TCAGAAGTCGAAAGCGCTGTAATTCGTTCACCACACCTATTTGATATACTCAGTGAAATA 1056

Qy        359 PheValTyrThrArgThrLeuGly---------AsnMetSerAspValArgAsnIleTrp 375
Db       1057 GAATTTTATACAACAAGAGCGGGGCTTCCCTTGAATAATACGGAATACCTTGAATATTGG 1116

Qy        376 GlyGlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPhe 394
Db       1117 GTAGGACATTCTATAAAATATAAAAATACGAATGCCTCATCAGCATTAGAACGTAATTAC 1176

Qy        395 GlyLysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPhe 414
Db       1177 GGT---ACGATTACTTCTAACAAAATCAAGTATTATGATTTAGCAAATAAGGATATCTTT 1233

Qy        415 SerIleGluSerLeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsnAsnTyrIleThr 434
Db       1234 CAGGTTCGATCATTA-----------------GGGGCGGATTTAGCTAATTACTACGCA 1275

Qy        435 SerGlnTyrGlyValSerArgValIlePheAsnThrSerAsnIleAsnAsnValProGly 454
Db       1276 CAGGTATATGGAGTTCCGTACGCTAGTTTTACACTGCTTGACAAGAATACAGGATCAGGA 1335

Qy        455 SerLeu---------ArgTyrGluValProAlaAsn---------------LeuProSer 466
Db       1336 TCAGTTGGAGGTTTTACGTACTCAAAACCACATACAACTATGCAAGTATGTACACAAAAT 1395

Qy        467 GlnThrIleLeuSerGluLeuProGlyLysAspLysProArgProAsnAlaGlyAspPhe 486
Db       1396 TACAATACGATTGATGAAATCCCTCCAGAGAATGAGCCACTTAGTAGAGGG------TAT 1449

Qy        487 SerHisArgLeuSerTyrIleSer------AsnPheAspAlaArgArgSerSerSerGly 504
Db       1450 AGCCATAGATTATCTCATATCACCTCTTATTCTTTTTCTAAGAATGCTAGTAGTCCTGCT 1509

Qy        505 GlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsnAsnArg 524
Db       1510 AGATATGGCAATCTCCCTGTATTTGCTTGGACACATCGGAGTGCGGATGTTACAAATACA 1569

Qy        525 LeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrp---------GlyGly 541
Db       1570 GTTTATTCAGATAAAATTACTCAGATACCAGTTGTAAAGGCACATACTTTAGTTTCAGGT 1629

Qy        542 AsnIleGlyPheValIleProGlyProThrGlyGlyAsnLeuValLysValSerAspSer 561
```

```
                              ||||||   ||||||||||||::::::|||    :::    |||
Db       1630 ACTACTGTTATTAAAGGTCCTGGATTTACAGGAGGCAATATCCTTAAAAGAACAAGTAGT 1689

Qy        562 ------------TrpHisSerLeuLysValGlnAlaProGlnArgGlnThrSerTyrArg 577
                          :::    |||:::    |||::::::|||      |||          ||||||
Db       1690 GGTCCGTTAGCTTATACTAGTGTCTCTGTAAAATCACCATTATCACAA---AGATATCGT 1746

Qy        578 IleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAlaIlePheValGluHisSer 597
                 |||:::||||||||            |||            :::||||||       |||
Db       1747 GCAAGAATACGTTATGCTTCT------ACTACTAACTTACGACTTTTTGTAACAATTTCT 1800

Qy        598 GlySerSerHisIleValSerPhePheAspCysSerAsnSerSerGlyArgProSerAsn 617
              |||   ||||||::::::        :::   |||   |||   :::              :::
Db       1801 GGAAC-TCGCATTTACTC------TATAAATGTT--AATAAAACCATGAATAAAGGGGAT 1851

Qy        618 ThrLeuLeuGluSerAspPheArgTyrIleAspValProGlyIlePheThr---ProSer 636
                 |||      :::   |||               :::         ||||||        :::
Db       1852 GATTTAACATTTAATACATTTGACTTAGCAACTATTGGTACTGCTTTCACATTTTCAAAT 1911

Qy        637 IleAsnProLeuIleArgTyrArgThrGlnSerPhe---------GlyThrHisAlaIle 653
                 :::      :::                  ||||||         |||              :::
Db       1912 TACTCGGATAGCTTAACGGTAGGTGCAGATTCTTTTGCTTCAGGAGGAGAAGTTTATGTA 1971

Qy        654 AspLysPheGluPheIleProLeuAsn---ThrPhe---ProAsnGlnSerLeuGluLys 671
              ||||||||||||   |||||||:::|||      ||||||         :::     |||:::
Db       1972 GATAAGTTCGAACTTATTCCGGTAAATGCAACATTTGAAGCAGAAGAAGACCTAGATGTG 2031

Qy        672 ArgGluGlnGluValAsnAspLeu 679
                 ::::::   |||   :::|||
Db       2032 GCAAAGAAAGCAGTAAAAAACTTA 2055
```

```
RESULT 14
AX543928
LOCUS       AX543928                2003 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 5 from Patent WO0234774.
ACCESSION   AX543928
VERSION     AX543928.1  GI:25277387
KEYWORDS    .
SOURCE      Bacillus thuringiensis
  ORGANISM  Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Abad,A.R., Duck,N.B., Feng,X., Flannagan,R.D., Kahn,T.W. and
            Sims,L.E.
  TITLE     Genes encoding novel proteins with pesticidal activity against
            coleopterans
  JOURNAL   Patent: WO 0234774-A 5 02-MAY-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1..2003
                     /organism="Bacillus thuringiensis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1428"
                     /note="(truncated)
                     1218-1"
     CDS             1..>2001
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD57544.1"
                     /db_xref="GI:25277388"
                     /translation="MSPNNQNEYEIIDATPSTSVSNDSNRYPFANEPTNALQNMDYKD
                     YLKMSAGNASEYPGSPEVLVSGQDAAKAAIDIVGKLLSGLGVPFVGPIVSLYTQLIDI
                     LWPSGEKSQWEIFMEQVEELINQKIAEYARNKALSELEGLGNNYQLYLTALEEWEENP
                     NGSRALRDVRNRFEILDSLFTQYMPSFRVTNFEVPFLTVYAMAANLHLLLLKDASIFG
                     EEWGWSTTTINNYYDRQMKLTAEYSDHCVKWYETGLAKLKGTSAKQWVDYNQFRREMT
                     LAVLDVVALFPNYDTRTYPMETKAQLTREVYTDPLGAVNVSSIGSWYDKAPSFGVIES
                     SVIRPPHVFDYITGLTVYTQSRSISSARYIRHWAGHQISYHRVSRGSNLQQMYGTNQN
                     LHSTSTFDFTNYDIYKTLSKDAVLLDIVYPGYTYIFFGMPEVEFFMVNQLNNTRKTLK
                     YNPVSKDIIASTRDSELELPPETSDQPNYESYSHRLCHITSIPATGNTTGLVPVFSWT
                     HRSADLNNTIYSDKITQIPAVKCWDNLPFVPVVKGPGHTGGDLLQYNRSTGSVGTLFL
                     ARYGLALEKAGKYRVRLRYATDADIVLHVNDAQIQMPKTMNPGEDLTSKTFKVADAIT
                     TLNLATDSSLALKHNLGEDPNSTLSGIVYVDRIEFIPVDETYE"
ORIGIN

Alignment Scores:
Pred. No.:              1.24e-59        Length:         2003
Score:                  926.00          Matches:        239
Percent Similarity:     50.8%           Conservative:   116
Best Local Similarity:  34.2%           Mismatches:     268
Query Match:            25.7%           Indels:         76
DB:                     6               Gaps:           23

US-10-782-096-2 (1-682) x AX543928 (1-2003)
```

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
              |||         |||:::||||||||||||||||:::::::||||||          ::::::::::::
Db          4 AGTCCAAATAATCAAAATGAATATGAAATTATAGATGCGACACCT---TCTACTTCTGTA 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||         |||||||||   |||   :::|||            ::::::|||
Db         61 TCCAATGATTCTAACAGATACCCTTTTGCGAATGAGCCAACAAATGCGCTACAAAATATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                           |||:::   |||
Db        121 GATTATAAAGATTATTTAAAAATGTCT------------------GCGGGAAATGCTAGT 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||         :::|||   :::||||||               |||
Db        163 GAATACCCTGGTTCACCTGAAGTACTTGTTAGCGGACAAGATGCAGCTAAGGCCGCAATT 222

Qy         82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
              :::      ||||||   :::||||||      ||||||||||||
Db        223 GATATAGTAGGTAAATTACTATCAGGTTTAGGGGTCCCATTTGTTGGGCCGATAGTGAGT 282

Qy        102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
                 :::   ::::::||||||   |||||||||                     :::|||
Db        283 CTTTATACTCAACTTATTGATATTCTGTGGCCTTCAGGGGAAAAGAGTCAATGGGAAATT 342

Qy        120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
                 |||      ||||||||||||||||   :::   |||   :::      |||
Db        343 TTTATGGAACAAGTAGAAGAACTCATTAATCAAAAAATAGCAGAATATGCAAGGAATAAA 402

Qy        140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
              ||||||   ||||||||||||||||               :::||||||   |||   |||:::
Db        403 GCGCTTTCGGAATTAGAAGGATTAGGTAATAATTACCAATTATATCTAACTGCGCTTGAA 462

Qy        160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
                 |||         |||   |||            |||      :::      ::::::
Db        463 GAATGGGAAGAAAATCCAAATGGTTCAAGAGCCTTACGAGATGTGCGAAATCGATTTGAA 522

Qy        179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
              |||:::|||:::      :::         ||||||   |||            |||||||||:::
Db        523 ATCCTGGATAGTTTATTTACGCAATATATGCCATCTTTTAGAGTGACAAATTTTGAAGTA 582

Qy        199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
                       |||   |||||||||   |||||||||||||||||:::||||||:::||||||
Db        583 CCATTCCTTACTGTATATGCAATGGCAGCCAACCTTCATTTACTGTTATTAAAGGACGCG 642

Qy        219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
                       ||||||   :::|||                     |||   :::   |||   |||
Db        643 TCAATTTTTGGAGAAGAATGGGGATGGTCAACAACTACTATTAATAACTATTATGATCGT 702

Qy        239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
                    |||      ||||||   |||||||||:::   :::|||         ||||||
Db        703 CAAATGAAACTTACTGCAGAATATTCTGATCACTGTGTAAAGTGGTATGAAACTGGTTTA 762

Qy        259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
              :::         ::::::|||:::      ||||||   :::|||:::|||||||      :::
Db        763 GCAAAATTAAAAGGCACGAGCGCTAAACAATGGGTTGACTATAACCAATTCCGTAGAGAA 822

Qy        279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
              |||||||||   |||||||||:::   ||||||||||||||||||   |||   |||
Db        823 ATGACACTGGCGGTTTTAGATGTTGTTGCATTATTCCCAAATTATGACACACGCACGTAC 882
```

```
Qy        299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
              |||:::      |||    :::|||||||||||||||||||||||||||:::|||
Db        883 CCAATGGAAACGAAAGCACAACTAACAAGGGAAGTATATACAGATCCACTGGGCGCGGTA 942

Qy        319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
                 |||         |||                 ||||||:::          :::|||
Db        943 AACGTGTCTTCAATTGGTTCC------------TGGTATGAC---AAAGCACCTTCTTTC 987

Qy        339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
                    :::|||::::::          |||   |||         ::::::      :::
Db        988 GGAGTGATAGAATCATCCGTTATTCGACCACCCCATGTATTGATTATATAACGGGACTC 1047

Qy        359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
                 ||||||||      |||::::::   :::         :::|||:::   |||
Db       1048 ACAGTGTATACACAATCAAGAAGCATTTCTTCCGCTCGCTATATAAGACAT---TGGGCT 1104

Qy        377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
              ||||||   :::                  :::   ||||||   :::          :::|||
Db       1105 GGTCATCAAATAAGCTACCATCGTGTCAGTAGGGGTAGTAATCTTCAACAAATGTATGGA 1164

Qy        396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
                    ::::::                |||:::|||   |||         ::::::
Db       1165 ACTAATCAAAATCTACACAGCACTAGTACCTTTGATTTTACGAATTATGATATTTACAAG 1224

Qy        416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
                    |||               |||      :::|||   |||   |||
Db       1225 ACTCTATCAAAGGATGCAGTACTCCTTGATATTGTTTACCCTGGTTATACG--------- 1275

Qy        431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
                 ||||||         :::|||:::      |||      |||      :::   :::
Db       1276 ---TATATATTT------TTTGGAATGCCAGAAGTCGAGTTTTTCATGGTAAACCAATTG 1326

Qy        450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
              ||||||         :::|||:::|||            |||      |||:::   |||
Db       1327 AATAATACCAGAAAGACGTTAAAGTAT------------AATCCAGTTTCCAAAGATATT 1374

Qy        470 LeuSer--------------------GluLeuProGlyLysAspLysProArgProAsn 482
              ::::::                    |||||||||   :::      :::||||||
Db       1375 ATAGCGAGTACAAGAGATTCGGAATTAGAATTACCTCCAGAAACTTCAGATCAACCAAAT 1434

Qy        483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
                    :::||||||||||||||   :::|||:::      |||         ::::::
Db       1435 TATGAGTCATATAGCCATAGATTATGTCATATCACAAGTATTCCCGCGACGGGTAACACT 1494

Qy        503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
              :::|||   :::                   |||   |||   |||   |||   |||   |||
Db       1495 ACCGGATTAGTACCT---------GTATTTTCTTGGACACATCGAAGTGCAGATTTAAAC 1545

Qy        523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
              |||   :::         |||||||||||||||   |||||||||   |||      |||
Db       1546 AATACAATATATTCAGATAAAATCACTCAAATTCCGGCCGTTAAATGTTGG---GATAAT 1602

Qy        543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
              :::   ||||||            ||||||   |||||||||:::|||::::::
Db       1603 TTACCGTTTGTTCCAGTGGTAAAAGGACCAGGACATACAGGAGGGGATTTATTACAGTAT 1662

Qy        559 SerAspSerTrpHisSerLeuLysVal--------------------GlnAlaProGln 571
              :::   |||   |||:::                    |||   :::
Db       1663 AATAGAAGTACTGGTTCTGTAGGAACCTTATTTCTAGCTCGATATGGCCTAGCATTAGAA 1722

Qy        572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
              :::            ||||||:::||||||||||||                |||||
Db       1723 AAAGCAGGGAAATATCGTGTAAGACTGAGATATGCTACT---------------GATGCA 1767

Qy        592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
                    |||   |||         :::   |||
Db       1768 GATATTGTATTGCATGTAAACGATGCTCAGATTCAG---------------ATGCCAAAA 1812

Qy        612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
              :::         |||         |||         |||:::         |||
Db       1813 ACAATGAACCCAGGTGAGGATCTGACATCTAAAACTTTTAAAGTTGCAGATGCTATCACA 1872

Qy        632 IlePheThrProSerIleAsnProLeuIleArgTyrArg-------------------- 644

                            :::   :::      :::      :::
Db       1873 ACATTAAATTTAGCAACAGATAGTTCGCTAGCATTGAAACATAATTTAGGTGAAGACCCT 1932

Qy        645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
                 :::         |||         :::|||:::   ||||||||||||::::::
Db       1933 AATTCAACATTATCTGGTATAGTTTACGTTGACCGAATCGAATTCATCCCAGTAGAT 1989
```

RESULT 15
CS130950

```
LOCUS       CS130950                2010 bp    DNA     linear   PAT 02-AUG-2005
DEFINITION  Sequence 5 from Patent WO2005066349.
ACCESSION   CS130950
VERSION     CS130950.1  GI:71793187
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Abad,A.
  TITLE     Genes encoding proteins with pesticidal activity
  JOURNAL   Patent: WO 2005066349-A 5 21-JUL-2005;
            E.I. Du pont de nemours and company (US); Pioneer Hi-Bred
            International, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..2010
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="mo1218-1"
     CDS             1..2010
                     /note="unnamed protein product; Maize optimized Cry1218-1"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAJ21053.1"
                     /db_xref="GI:71793188"
                     /translation="MSPNNQNEYEIIDATPSTSVSNDSNRYPFANEPTNALQNMDYKD
                     YLKMSAGNASEYPGSPEVLVSGQDAAKAAIDIVGKLLSGLGVPFVGPIVSLYTQLIDI
                     LWPSGEKSQWEIFMEQVEELINQKIAEYARNKALSELEGLGNNYQLYLTALEEWEENP
                     NGSRALRDVRNRFEILDSLFTQYMPSFRVTNFEVPFLTVYAMAANLHLLLLKDASIFG
                     EEWGWSTTTINNYYDRQMKLTAEYSDHCVKWYETGLAKLKGTSAKQWVDYNQFRREMT
                     LAVLDVVALFPNYDTRTYPMETKAQLTREVYTDPLGAVNVSSIGSWYDKAPSFGVIES
                     SVIRPPHVFDYITGLTVYTQSRSISSARYIRHWAGHQISYHRVSRGSNLQQMYGTNQN
                     LHSTSTFDFTNYDIYKTLSKDAVLLDIVYPGYTYIFFGMPEVEFFMVNQLNNTRKTLK
                     YNPVSKDIIASTRDSELELPPETSDQPNYESYSHRLCHITSIPATGNTTGLVPVFSWT
                     HRSADLNNTIYSDKITQIPAVKCWDNLPFVPVVKGPGHTGGDLLQYNRSTGSVGTLFL
                     ARYGLALEKAGKYRVRLRYATDADIVLHVNDAQIQMPKTMNPGEDLTSKTFKVADAIT
                     TLNLATDSSLALKHNLGEDPNSTLSGIVYVDRIEFIPVDETYEAE"
ORIGIN
```

Alignment Scores:

Pred. No.:	1.25e-59	Length:	2010
Score:	926.00	Matches:	239
Percent Similarity:	50.8%	Conservative:	116
Best Local Similarity:	34.2%	Mismatches:	268
Query Match:	25.7%	Indels:	76
DB:	6	Gaps:	23

US-10-782-096-2 (1-682) x CS130950 (1-2010)

```
Qy          3 SerTyrLysAsnLysAsnGluTyrGluMetLeuAspAlaLeuArgIleAsnSerAsnMet 22
                 |||         |||:::||||||||||||::::::||||||         ::::::::::::
Db          4 TCCCCCAACAACCAGAACGAGTACGAGATCATCGACGCCACCCCC---TCCACCTCCGTG 60

Qy         23 SerAsnCysTyrProArgTyrProLeuAlaLysAspProGlnMetThrMetArgAsnThr 42
              ||||||            |||||||||   |||   :::|||         ::::::|||
Db         61 TCCAACGACTCCAACCGCTACCCCTTCGCCAACGAGCCCACCAACGCCCTCCAGAACATG 120

Qy         43 AsnTyrLysGluTrpLeuAsnMetCysAspSerAsnThrGlnPheIleGlyAspIleSer 62
              :::||||||::::::|||   |||                        |||:::   |||
Db        121 GACTACAAGGACTACCTCAAGATGTCC------------------GCCGGCAACGCCTCC 162

Qy         63 ThrTyr---SerSerProGluAlaAlaLeuSerValArgAspAlaValLeuThrGlyIle 81
                 |||      |||||||||      :::|||   :::||||||         |||
```

```
Db    163 GAGTACCCCGGCTCCCCCGAGGTGCTCGTGTCCGGCCAGGACGCCGCCAAGGCCGCCATC 222

Qy     82 AsnSerValGlyThrIleLeuSerAsnLeuGlyValProLeuAlaSerGlnSerPheGly 101
Db    223 GACATCGTGGGCAAGCTCCTCTCCGGCCTCGGCGTGCCCTTCGTGGGCCCCATCGTGTCC 282

Qy    102 IleIleSerArgLeuIleGlyIleLeuTrpAlaGlyProAsp------ProPheGluAla 119
Db    283 CTCTACACCCAGCTCATCGACATCCTCTGGCCCTCCGGCGAGAAGTCCCAGTGGGAAATC 342

Qy    120 LeuMetValLeuValGluGluLeuIleLysLysSerIleAspGlnArgValArgGluAsn 139
Db    343 TTCATGGAGCAGGTGGAGGAGCTCATCAACCAGAAGATCGCCGAGTACGCCCGCAACAAG 402

Qy    140 AlaLeuArgGluLeuGluGlyLeuGlnGlyIleMetArgLeuTyrGlnThrArgLeuGln 159
Db    403 GCCCTCTCCGAGCTGGAGGGCCTCGGCAACAACTACCAGCTCTACCTCACCGCCCTGGAG 462

Qy    160 AlaTrpLeuValAsnLysAsnAspAspAsnArg---ArgAlaLeuValThrGlnTyrAla 178
Db    463 GAGTGGGAGGAGAACCCCAACGGCTCCCGCGCCCTCCGCGACGTGCGCAACCGCTTCGAG 522

Qy    179 IleValAspAsnPhePheGluLysAsnMetProLysPheLysGluArgAsnPheGluIle 198
Db    523 ATCCTCGACTCCCTCTTCACCCAGTACATGCCCTCCTTCCGCGTGACCAACTTCGAGGTG 582

Qy    199 LeuLeuLeuProValTyrAlaGlnAlaAlaAsnLeuHisLeuIleLeuLeuArgAspAla 218
Db    583 CCCTTCCTCACCGTGTACGCCATGGCCGCCAACCTCCACCTCCTCCTCCTCAAGGACGCC 642

Qy    219 AspTyrPheGlyAlaGlnTrpGlnLeuGlyAspAspGluIleArgAspAsnTyrIleArg 238
Db    643 TCCATCTTCGGCGAGGAGTGGGGCTGGTCCACCACCACCATCAACAACTACTACGACCGC 702

Qy    239 LeuGlnGlyLeuIleArgGluTyrLysAspHisCysIleThrPheTyrAsnGlnGlyLeu 258
Db    703 CAGATGAAGCTCACCGCCGAGTACTCCGACCACTGCGTGAAGTGGTATGAGACCGGCCTC 762

Qy    259 AsnGlnPheAsnArgSerAsnAlaGlnAspTrpValSerPheAsnArgPheArgThrAsp 278
Db    763 GCCAAGCTCAAGGGCACCTCCGCCAAGCAGTGGGTGGACTACAACCAGTTCCGCCGCGAG 822

Qy    279 MetThrLeuThrValLeuAspLeuAlaIleLeuPheProAsnTyrAspProArgArgTyr 298
Db    823 ATGACCCTCGCCGTGCTCGACGTGGTGGCCCTCTTCCCCAACTACGACACCCGCACCTAC 882

Qy    299 ProLeuAlaValLysThrGluLeuThrArgGluValTyrThrAspProValGlyPheThr 318
Db    883 CCCATGGAGACCAAGGCCCAGCTCACCCGCGAGGTGTACACCGACCCGCTCGGCGCCGTG 942

Qy    319 GlyValLeuGluSerGlyGlyArgThrTyrProTrpTyrAsnProAsnAsnThrThrPhe 338
Db    943 AACGTGTCCTCCATCGGCTCT------------TGGTACGAC---AAGGCCCCAAGCTTC 987

Qy    339 ThrAlaMetGluAsnAsnAlaArgArgArgProSerTyrThrThrTrpLeuAsnArgIle 358
Db    988 GGCGTGATCGAGTCCTCCGTGATCCGCCCGCCGCACGTGTTCGACTACATCACCGGCCTC 1047

Qy    359 PheValTyrThr------ArgThrLeuGlyAsnMetSerAspValArgAsnIleTrpGly 376
Db   1048 ACCGTGTACACCCAGTCCCGCTCCATCTCCTCCGCCCGCTACATCCGCCAC---TGGGCC 1104

Qy    377 GlyHisThrLeu---ValGluAsnGlyAsnAspGlySerGluIleThrHisAsnPheGly 395
Db   1105 GGCCACCAGATCTCCTACCACCGCGTGTCCCGCGGCTCCAACCTCCAGCAGATGTACGGC 1164

Qy    396 LysThrAspSerIleThrProIleGlnTyrPheAsnPheAlaAsnLeuSerValPheSer 415
Db   1165 ACCAACCAGAACCTCCACTCCACCTCCACCTTCGACTTCACCAACTACGACATCTACAAG 1224

Qy    416 IleGluSer---------------LeuAlaArgIleTyrLeuGlyGlyThrGluAlaAsn 430
Db   1225 ACCCTCTCCAAGGACGCCGTGCTCCTCGACATCGTGTACCCCGGCTACACC--------- 1275

Qy    431 AsnTyrIleThrSerGlnTyrGlyValSerArgVal---IlePheAsnThrSerAsnIle 449
Db   1276 ---TACATCTTC------TTCGGCATGCCGGAGGTGGAGTTCTTCATGGTGAACCAGCTC 1326

Qy    450 AsnAsnValProGlySerLeuArgTyrGluValProAlaAsnLeuProSerGlnThrIle 469
Db   1327 AACAACACCCGCAAGACCCTCAAATAC------------AACCCCGTGTCCAAGGACATC 1374

Qy    470 LeuSer---------------------GluLeuProGlyLysAspLysProArgProAsn 482
Db   1375 ATCGCCTCCACCCGCGACTCCGAGCTCGAGCTCCCCCCCGAGACCTCCGACCAGCCCAAC 1434

Qy    483 AlaGlyAspPheSerHisArgLeuSerTyrIleSerAsnPheAspAlaArgArgSerSer 502
Db   1435 TACGAGTCCTACTCCCACCGCCTCTGCCACATCACCTCCATCCCCGCCACCGGCAACACC 1494

Qy    503 SerGlyGlyIleValSerLeuLeuThrPheGlyTrpAlaHisThrSerMetAspArgAsn 522
Db   1495 ACCGGCCTCGTGCCG---------GTGTTCTCCTGGACCCACCGCTCTGCAGACCTCAAC 1545

Qy    523 AsnArgLeuGluProAspLysIleThrGlnIleAspAlaValLysGlyTrpGlyGlyAsn 542
Db   1546 AACACCATCTACTCCGACAAGATCACCCAGATCCCCGCCGTGAAGTGCTGG---GACAAC 1602

Qy    543 IleGlyPheValIle------------ProGlyProThrGlyGlyAsnLeuValLysVal 558
Db   1603 CTCCCCTTCGTGCCCGTGGTGAAGGGCCCCGGCCACACCGGCGGCGACCTCCTCCAGTAC 1662

Qy    559 SerAspSerTrpHisSerLeuLysVal---------------------GlnAlaProGln 571
Db   1663 AACCGCTCCACCGGCTCCGTGGGCACCCTCTTCCTCGCCCGCTACGGCCTCGCCCTGGAG 1722

Qy    572 ArgGlnThrSerTyrArgIleArgLeuArgTyrAlaCysLeuValThrHisGlyAspAla 591
Db   1723 AAGGCCGGCAAGTACCGCGTGCGCCTCCGCTACGCCACT---------------GACGCC 1767

Qy    592 IlePheValGluHisSerGlySerSerHisIleValSerPhePheAspCysSerAsnSer 611
Db   1768 GACATCGTGCTCCACGTGAACGACGCCCAGATCCAG---------------ATGCCCAAG 1812

Qy    612 SerGlyArgProSerAsnThrLeuLeuGluSerAspPheArgTyrIleAspValProGly 631
Db   1813 ACCATGAACCCCGGCGAGGACCTCACCTCCAAGACCTTCAAGGTGGCCGACGCCATCACC 1872

Qy    632 IlePheThrProSerIleAsnProLeuIleArgTyrArg--------------------- 644
Db   1873 ACCCTCAACCTCGCCACCGACTCCTCCCTCGCCCTCAAGCACAACCTCGGCGAGGACCCC 1932

Qy    645 ---ThrGlnSerPheGlyThrHisAlaIleAspLysPheGluPheIleProLeuAsn 662
Db   1933 AACTCCACCCTCTCCGGCATCGTGTACGTGGACCGCATCGAGTTCATCCCCGTGGAC 1989
```

Search completed: January 21, 2006, 01:57:09
Job time : 9422 secs